



University  
of Glasgow

Devlin, Rebecca (2015) *An investigation into the initiation of VSG switching in Trypanosoma brucei*. PhD thesis.

<http://theses.gla.ac.uk/6987/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

An investigation into the initiation of  
VSG switching in  
*Trypanosoma brucei*

**Rebecca Devlin**

BSc (Hons) MRes

Submitted in fulfilment of the requirements for the  
Degree of Doctor of Philosophy

Wellcome Trust Centre for Molecular Parasitology  
Institute of Infection, Immunity and Inflammation  
College of Medical Veterinary and Life Sciences  
University of Glasgow

December 2015



## Abstract

*Trypanosoma brucei*, the eukaryotic parasite that causes human African trypanosomiasis in humans, evades the immune system through antigenic variation. *T. brucei* antigenic variation involves the periodic switching of the variant surface glycoprotein (VSG) coat to an antigenically distinct variant. A single VSG is expressed on the cell surface at any one time, but the *T. brucei* genome contains a vast number of silent VSGs. To be expressed, a VSG must be located in a specialised VSG blood stream form expression site (VSG BES). Silent VSGs are copied into VSG BES by homologous recombination. Several proteins have been demonstrated to be involved in this process but how VSG switching is initiated remains unclear.

Four putative DNA repair factors were identified in *T. brucei*, whose eukaryotic homologues play a range of roles in DNA repair and other aspects of genome maintenance. These were two RecQ-like helicases, a Mus81 endonuclease and a Pif1 family helicase (PIF6). To examine whether these factors play a role in DNA repair and VSG switching, mutants were generated in blood stream form *T. brucei* cells. Analysis of RecQ1 by RNAi knockdown revealed it to be an essential gene in bloodstream form *T. brucei*, possibly involved in nuclear DNA replication. Phenotypic analysis of *recq2* mutants suggests that RECQ2 is involved in the repair of a range of DNA damaging agents. Furthermore, analysis of survival following DSB induction suggests RECQ2 is involved in the repair of DNA DSBs, including those in the active VSG BES. VSG switching analysis showed that *recq2*<sup>-/-</sup> mutants have an elevated VSG switching rate and increase in recombination events upstream of the active VSG. These analyses suggest that RECQ2 suppresses VSG switching in *T. brucei* by suppressing recombination events near the active VSG. Analysis of *mus81* mutants showed *mus81*<sup>-/-</sup> mutants to be sensitive to agents inducing replication stalling and DNA breaks, and that MUS81 is important in the repair of DSBs. PIF6 appears to be a complicated DNA repair factor, different from MUS81 and RECQ2. *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> mutants appear to be more resistant to MMS than wild type cells, though more sensitive to the replication stalling agent hydroxyurea. *pif6* mutants do not appear to be more sensitive to DSBs than wild type cells and may even be more resistant. It is unclear whether PIF6 is involved in VSG switching and more work is required on

this factor to attempt to understand its DNA repair and VSG switching function in *T. brucei*.

These analyses shed light on the DNA repair functions of four previously uncharacterised *T. brucei* proteins. In particular, observations that RECQ2 is deficient in repairing DSBs upstream of the active VSG and mutants exhibit an elevated VSG switching rate cannot be reconciled with current thinking that direct formation of DSBs in this location initiates VSG switching. This suggests that the initiation of VSG switching is more complex than currently thought and requires careful further study and consideration of the relevance of using direct DSBs in this location to model VSG switching.

# Table of contents

|  |    |
|--|----|
| Abstract.....  | 2  |
| List of Tables .....   | 10 |
| List of Figures .....  | 11 |
| Acknowledgements .....   | 14 |
| Author's declaration .....   | 16 |
| Definitions and Abbreviations .....  | 17 |
| Chapter 1  |    |
| Introduction.....  | 21 |
| 1.1 General Introduction .....   | 22 |
| 1.1.1 <i>Trypanosoma brucei</i> life cycle .....   | 23 |
| 1.1.2 Human African Trypanosomiasis: prevalence, symptoms and treatment.....                   | 25 |
| 1.1.3 <i>T. brucei</i> genome .....  | 26 |
| 1.2 <i>T. brucei</i> antigenic variation .....   | 28 |
| 1.2.1 <i>T. brucei</i> variant surface glycoproteins (VSGs) .....                              | 29 |
| 1.2.2 VSG expression sites .....   | 31 |
| 1.2.3 Mechanisms of VSG switching in <i>T. brucei</i> .....                                    | 32 |
| 1.2.3.1 Transcriptional VSG switching.....   | 33 |
| 1.2.3.2 Recombinational VSG switching.....   | 34 |
| 1.2.4 <i>T. brucei</i> 70 bp repeats .....   | 38 |
| 1.2.4.1 DNA double strand breaks at 70 bp repeats as initiators of VSG switching.....          | 38 |
| 1.2.4.2 70 bp instability .....  | 42 |
| 1.3 Double Strand Break Repair (DSBR).....   | 44 |
| 1.3.1 Non-homologous end-joining (NHEJ) .....  | 45 |
| 1.3.2 Homologous recombination (HR).....   | 46 |
| 1.3.2.1 Gene conversion.....   | 48 |
| 1.3.2.2 Break-induced replication (BIR).....   | 49 |
| 1.3.2.3 Synthesis-dependent strand annealing (SDSA).....                                       | 49 |
| 1.3.3 Single strand annealing (SSA).....   | 50 |
| 1.3.4 Microhomology-mediated end joining (MMEJ) .....  | 51 |
| 1.4 DNA Helicases in DNA repair .....  | 52 |
| 1.4.1 RecQ helicases in DNA repair and two putative <i>T. brucei</i> RecQ-like helicases ..... | 54 |
| 1.4.2 Pif1 family helicases in DNA repair .....  | 58 |
| 1.4.2.1 Pif1 helicases in <i>T. brucei</i> .....   | 61 |
| 1.5 Mus81 endonucleases in DNA repair and a putative <i>T. brucei</i> Mus81 homologue .....    | 61 |
| 1.6 Aims of this project .....   | 63 |

|       |  |    |
|-------|--|----|
| 1.6.1 | Roles of four factors in DNA repair .....                          | 63 |
| 1.6.2 | Roles of four factors in DSB repair.....                           | 64 |
| 1.6.3 | Roles of four factors in VSG switching .....                       | 64 |
| 1.6.4 | DSBs at the 70 bp repeats as the initiators of VSG switching ..... | 64 |

## Chapter 2

|  |           |
|--|-----------|
| <b>Materials and Methods .....</b>   | <b>65</b> |
| 2.1 Trypanosome culture .....  | 66        |
| 2.1.1 Trypanosome strains and their growth.....  | 66        |
| 2.1.2 Stabilate preparation and retrieval .....  | 67        |
| 2.1.3 Transfection of trypanosomes.....  | 67        |
| 2.1.4 Preparation of trypanosomes for molecular biology procedures.....                        | 68        |
| 2.2 Analysis of trypanosome phenotypes .....   | 68        |
| 2.2.1 Cell cycle analysis.....   | 68        |
| 2.2.2 Analysis of trypanosome in vitro growth.....   | 69        |
| 2.2.3 Analysis of DNA damage sensitivity.....  | 69        |
| 2.2.3.1 DNA damage growth curve analysis .....   | 69        |
| 2.2.3.2 DNA damage clonal survival assays .....  | 70        |
| 2.2.3.3 I-SceI double strand break survival assays .....                                       | 70        |
| 2.2.4 VSG switching analysis.....  | 71        |
| 2.2.4.1 Preparation of media for the culture of VSG switching cell<br><i>GFP221hygTK</i> ..... | 71        |
| 2.2.4.2 Analysis of VSG switching frequency and mechanism .....                                | 71        |
| 2.3 Trypanosome RNA interference (RNAi).....   | 73        |
| 2.3.1 Cloning of <i>RECQ1</i> RNAi plasmid using Gateway® system.....                          | 73        |
| 2.4 Isolation of material from trypanosomes .....  | 74        |
| 2.4.1 Isolation of genomic DNA .....   | 74        |
| 2.4.2 Isolation of total RNA.....  | 74        |
| 2.5 Electrophoresis .....  | 74        |
| 2.5.1 DNA electrophoresis .....  | 74        |
| 2.5.2 Protein electrophoresis.....   | 74        |
| 2.6 Blotting.....  | 75        |
| 2.6.1 Western blot transfer.....   | 75        |
| 2.6.2 Western blot detection.....  | 75        |
| 2.6.3 Hybridisation and detection of antibodies .....  | 75        |
| 2.6.4 Stripping western blots.....   | 76        |
| 2.7 Molecular biology techniques.....  | 78        |
| 2.7.1 Polymerase chain reaction (PCR) .....  | 78        |
| 2.7.2 DNA fragment gel extraction .....  | 78        |
| 2.7.3 cDNA synthesis.....  | 78        |
| 2.7.4 Restriction enzyme analysis of DNA .....   | 79        |

|         |  |    |
|---------|--|----|
| 2.7.5   | Primer design .....  | 79 |
| 2.7.6   | Plasmid sequencing .....                                     | 79 |
| 2.7.7   | Concentration of digested DNA .....                          | 79 |
| 2.8     | Cloning .....  | 79 |
| 2.8.1   | T4 DNA ligase .....  | 79 |
| 2.9     | Transformation of <i>E. coli</i> and plasmid retrieval ..... | 80 |
| 2.9.1   | Transformation of <i>E. coli</i> .....                       | 80 |
| 2.9.1.1 | Preparation of ZYM-5 medium .....                            | 80 |
| 2.9.2   | Extraction of plasmid DNA from <i>E. coli</i> .....          | 81 |
| 2.10    | Microscopy .....   | 81 |
| 2.10.1  | DAPI staining.....   | 81 |
| 2.10.2  | Immunofluorescence microscopy.....                           | 82 |
| 2.11    | Bioinformatics .....   | 82 |

### Chapter 3

|   |   |           |
|---|---|-----------|
| <b>Analysis of the roles in general DNA repair of RECQ2, PIF6 and MUS81 .....</b> |   | <b>84</b> |
| 3.1   | Introduction .....  | 85        |
| 3.2   | Sequence analysis of putative DNA repair factors.....               | 85        |
| 3.2.1   | <i>T. brucei</i> RECQ2 is a RecQ helicase homologue .....           | 85        |
| 3.2.2   | Sequence analysis of <i>T. brucei</i> MUS81 .....                   | 91        |
| 3.2.3   | TbPIF6 is a member of the PIF1-like helicase family .....           | 95        |
| 3.3   | Generation of knockout mutants.....                                 | 100       |
| 3.3.1   | Knockout strategy .....   | 100       |
| 3.3.2   | Generation of knockout mutants in the Lister 427 cell line .....    | 102       |
| 3.3.3   | Confirmation of knockout mutants by PCR.....                        | 103       |
| 3.3.4   | Confirmation of knockout mutants by RT-PCR.....                     | 105       |
| 3.4   | Phenotypic analysis of knockout mutants .....                       | 106       |
| 3.4.1   | Analysis of <i>in vitro</i> growth .....                            | 106       |
| 3.4.2   | Cell cycle analysis.....  | 108       |
| 3.4.3   | Analysis of DNA damage sensitivity.....                             | 109       |
| 3.4.3.1   | Analysis of <i>in vitro</i> growth after DNA damage.....            | 110       |
| 3.4.3.2   | Analysis of clonal survival after DNA damage.....                   | 113       |
| 3.5   | Generation of re-expresser cell lines.....                          | 119       |
| 3.5.1   | Generation of re-expresser constructs.....                          | 119       |
| 3.5.2   | Confirmation of re-expresser cell lines by PCR and RT-PCR.....      | 120       |
| 3.6   | Analysis of re-expresser cell lines .....                           | 122       |
| 3.6.1   | Analysis of <i>in vitro</i> growth of re-expresser cell lines ..... | 122       |
| 3.6.2   | Analysis of DNA damage sensitivity of re-expresser cell lines.....  | 123       |
| 3.7   | Generation of endogenously tagged cell lines .....                  | 124       |
| 3.7.1   | Generation of C-terminal myc-tagged MUS81 and PIF6 cell lines...    | 124       |
| 3.7.2   | Generation of N-terminal myc-tagged RECQ2 cell line.....            | 126       |

|          |   |     |
|----------|---|-----|
| 3.7.3    | Confirmation of <i>in vivo</i> myc-tagging by western blot analysis ..... | 129 |
| 3.8      | Testing the functionality of endogenously-tagged proteins .....           | 130 |
| 3.8.1    | Generation and validation of single-copy myc tagged cell lines....        | 130 |
| 3.8.2    | Testing functionality using DNA damage sensitivity assays .....           | 133 |
| 3.9      | Localisation of endogenously-tagged factors .....                         | 134 |
| 3.9.1    | Localisation of 12myc-RECQ2.....  | 135 |
| 3.9.2    | Localisation of MUS81myc.....   | 138 |
| 3.9.3    | Localisation of PIF6myc .....   | 140 |
| 3.10     | Summary.....  | 143 |
| 3.10.1   | Key findings of this chapter .....  | 143 |
| 3.10.1.1 | <i>T. brucei</i> RECQ2.....   | 143 |
| 3.10.1.2 | <i>T. brucei</i> MUS81.....   | 145 |
| 3.10.1.3 | <i>T. brucei</i> PIF6 .....   | 146 |

## Chapter 4

### Analysis of the role in double strand break (DSB) DNA repair and VSG

|  |     |
|--|-----|
| switching of RECQ2, PIF6 and MUS81 .....   | 149 |
| 4.1 Introduction .....   | 150 |
| 4.2 Description of I-SceI cell lines used .....  | 152 |
| 4.3 Generation of knockout mutants in I-SceI lines.....  | 154 |
| 4.3.1 Generation of knockout lines.....  | 154 |
| 4.3.2 Confirmation of knockout mutants by PCR.....   | 155 |
| 4.3.3 Confirmation of <i>recq2</i> and <i>mus81</i> knockout lines by RT-PCR.....                                      | 159 |
| 4.4 Analysis of DSB repair at a chromosome internal location.....  | 161 |
| 4.4.1 Analysis of chromosome-internal DSB repair in <i>recq2</i> mutants .....   | 163 |
| 4.4.2 Analysis of chromosome-internal DSB repair in <i>mus81</i> mutants ....  | 167 |
| 4.4.3 Analysis of chromosome-internal DSB repair in <i>pif6</i> mutants .....  | 169 |
| 4.5 Analysis of DSB repair in the active VSG expression site .....   | 171 |
| 4.5.1 Analysis of active VSG expression site DSB repair in <i>recq2</i> mutants  | 173 |
| 4.5.2 Analysis of active VSG expression site DSB repair in <i>mus81</i> mutants ..                                     | 177 |
| 4.5.3 Analysis of active expression site DSB repair in <i>pif6</i> mutants.....  | 180 |
| 4.6 Analysis of cell cycle, VSG221 expression and H2Ax nuclear signal after DSB induction in <i>pif6</i> mutants ..... | 185 |
| 4.6.1 Analysis of cell cycle arrest following DSB induction in <i>HRES</i> and <i>HR1 pif6</i> mutants .....           | 185 |
| 4.6.2 Analysis of VSG221 expression after DSB induction in <i>HRES pif6</i> mutants .....                              | 189 |
| 4.6.3 Analysis of H2Ax nuclear signal after DSB induction in <i>HRES</i> and <i>HR1 pif6</i> mutants .....             | 192 |
| 4.7 Analysis of VSG switching in knockout cell lines.....  | 199 |
| 4.7.1 VSG switching assay strategy .....   | 199 |

|         |  |     |
|---------|--|-----|
| 4.7.2   | Generation of VSG switching cell line.....                             | 203 |
| 4.7.3   | Generation and confirmation of mutants in VSG switching cell line..... | 207 |
| 4.7.4   | VSG switching analysis.....  | 212 |
| 4.7.4.1 | Analysis of VSG switching in <i>recq2</i> mutants.....                 | 214 |
| 4.7.4.2 | Analysis of VSG switching in <i>pif6</i> mutants.....                  | 231 |
| 4.8     | Summary .....  | 241 |
| 4.8.1   | RECQ2 is involved in DSB repair and suppresses VSG switching.....      | 241 |
| 4.8.2   | <i>mus81</i> mutants are deficient in DSB repair .....                 | 245 |
| 4.8.3   | PIF6: a complicated factor in <i>T. brucei</i> nuclear DNA repair..... | 246 |

## Chapter 5

|  |  |     |
|--|--|-----|
| Analysis of the role of <i>T. brucei</i> RECQ1 in DNA repair ..... |  | 249 |
| 5.1  | Introduction .....   | 250 |
| 5.2  | Sequence analysis of TbRECQ1 .....                                     | 250 |
| 5.2.1  | Identification of TbRECQ1 .....  | 250 |
| 5.2.2  | Phylogenetic analysis .....  | 255 |
| 5.3  | Generation of a RECQ1 RNAi cell line .....                             | 256 |
| 5.3.1  | Attempts at generating a RECQ1 knockout cell line .....                | 256 |
| 5.3.2  | Generation of a RECQ1 RNAi construct .....                             | 261 |
| 5.3.2.1  | RNAi in <i>T. brucei</i> .....   | 261 |
| 5.3.2.2  | Gateway RNAi system and cloning of <i>RECQ1</i> RNAi construct ...     | 261 |
| 5.3.3  | Generation of <i>RECQ1</i> RNAi cell line .....                        | 266 |
| 5.3.4  | Confirmation of RECQ1 RNAi cell line by PCR.....                       | 267 |
| 5.4  | Phenotypic analysis of RECQ1 RNAi.....                                 | 268 |
| 5.4.1  | Effect of RECQ1 RNAi knockdown on growth.....                          | 268 |
| 5.4.2  | Analysis of cell cycle.....  | 268 |
| 5.4.3  | Analysis of methyl methanesulfonate damage sensitivity after RNAi..... | 272 |
| 5.4.4  | Analysis of $\gamma$ H2A intensity following <i>RECQ1</i> RNAi .....   | 273 |
| 5.4.5  | Confirmation of RECQ1 knockdown by western blot analysis .....         | 277 |
| 5.5  | Immunolocalisation of RECQ1 .....                                      | 280 |
| 5.5.1  | Generation of an endogenously myc-tagged RECQ1 cell line.....          | 280 |
| 5.5.2  | Confirmation of myc-tagged line by western blot analysis .....         | 280 |
| 5.5.3  | Testing the functionality of endogenously-tagged RECQ1 .....           | 281 |
| 5.5.4  | Localisation of RECQ1myc .....   | 282 |
| 5.6  | Summary .....  | 284 |
| 5.6.1  | TbRECQ1 is a diverged RecQ-like helicase .....                         | 284 |
| 5.6.2  | TbRECQ1 is essential in bloodstream form <i>T. brucei</i> .....        | 285 |
| 5.6.3  | RNAi suggests TbRECQ1 has a nuclear function.....                      | 285 |

## Chapter 6

|                  |                    |     |
|------------------|--------------------|-----|
| Discussion ..... |                    | 288 |
| 6.1              | Introduction ..... | 289 |

|                                 |   |            |
|---------------------------------|---|------------|
| 6.2                             | <i>TbRECQ2</i> is a DNA repair helicase that suppresses VSG switching .....                                     | 290        |
| 6.2.1                           | Summary .....   | 294        |
| 6.3                             | <i>TbMUS81</i> is a DNA repair factor acting in DSB repair .....  | 295        |
| 6.3.1                           | Summary .....   | 297        |
| 6.4                             | <i>TbPIF6</i> is an enigmatic DNA repair factor.....  | 297        |
| 6.4.1                           | Summary .....   | 299        |
| 6.5                             | <i>TbRECQ1</i> is an essential RecQ helicase in blood stream form cells, with a possible nuclear function ..... | 300        |
| 6.5.1                           | Summary .....   | 303        |
| 6.6                             | Conclusion .....  | 303        |
| <b>Appendices .....</b>         |   | <b>305</b> |
| 7.1                             | Oligonucleotide sequences.....  | 306        |
| 7.2                             | Protein accession numbers.....  | 308        |
| 7.3                             | Individual <i>pif6</i> mutant MMS clonal survival assays .....  | 310        |
| 7.4                             | HRES <i>recq2</i> I-SceI assay PCRs .....   | 312        |
| 7.5                             | HRES <i>mus81</i> I-SceI assay PCRs .....   | 316        |
| 7.6                             | HRES <i>pif6</i> I-SceI assay PCR .....   | 319        |
| <b>List of References .....</b> |   | <b>328</b> |



## List of Tables

|           |  |     |
|-----------|--|-----|
| Table 2-1 | List of antisera used in this thesis .....   | 77  |
| Table 4-1 | Survival data from first <i>RECQ2</i> VSG switching assay .....  | 216 |
| Table 4-2 | Summary of the analysis of clones from first <i>RECQ2</i> VSG switching experiment. ....                           | 219 |
| Table 4-3 | Survival in a second <i>RECQ2</i> VSG switching assay .....  | 220 |
| Table 4-4 | Summary of analysis of clones from a second <i>RECQ2</i> VSG switching experiment. ....                            | 228 |
| Table 4-5 | Number of survivors and survival rate in <i>PIF6</i> VSG switching assay   | 232 |
| Table 4-6 | Summary of the analysis of clones from <i>PIF6</i> VSG switching experiment. ....                                  | 240 |
| Table 5-1 | Top protein TbRECQ1 BLAST hits.....  | 253 |
| Table 6-1 | Summary of the data presented <i>recq2</i> , <i>mus81</i> and <i>pif6</i> mutants and <i>recq1</i> RNAi cells..... | 290 |
| Table 7-1 | List of oligonucleotides used in this thesis .....   | 306 |
| Table 7-2 | Accession numbers for sequences used in sequence analysis.....   | 308 |
| Table 7-3 | Summary of analysis of first <i>RECQ2</i> HRES I-SceI assay .....  | 315 |
| Table 7-4 | Summary of analysis of first <i>RECQ2</i> HRES I-SceI assay .....  | 318 |
| Table 7-5 | Summary of analysis of survivors from first <i>PIF6</i> HRES I-SceI assay.   | 323 |

## List of Figures

|             |   |     |
|-------------|---|-----|
| Figure 1-1  | <i>T. brucei</i> life cycle .....   | 24  |
| Figure 1-2  | The VSG surface coat of <i>T. brucei</i> .....  | 30  |
| Figure 1-3  | Generalised schematic of a bloodstream VSG BES.....   | 32  |
| Figure 1-4  | Transcriptional VSG switching .....   | 34  |
| Figure 1-5  | Pathways of VSG switching by recombination.....   | 37  |
| Figure 1-6  | Effect of DSB induction at the 70 bp repeats on VSG switching. ....   | 41  |
| Figure 1-7  | Schematic of the steps of non-homologous end joining .....  | 46  |
| Figure 1-8  | Schematic of homologous recombination in eukaryotes.....  | 48  |
| Figure 1-9  | Pathways of DSB repair .....  | 50  |
| Figure 1-10 | DSB repair by single strand annealing (SSR).....  | 51  |
| Figure 1-11 | Microhomology-mediated end joining (MMEJ) .....   | 52  |
| Figure 1-12 | Domains and features of bacterial and eukaryotic RecQ helicases   | 54  |
| Figure 1-13 | Resolution and dissolution of double Holliday junctions .....   | 57  |
| Figure 1-14 | Pif1 family helicases .....   | 59  |
| Figure 1-15 | Illustration of yeast and human Mus81 proteins .....  | 63  |
| Figure 2-1  | Formulae used to calculate the number of generations in VSG switching experiments .....                     | 72  |
| Figure 3-1  | Representation of predicted protein domains in TbRECQ2.....   | 87  |
| Figure 3-2  | Multiple sequence alignment of RecQ-like helicases.....   | 88  |
| Figure 3-3  | Neighbour joining tree of eukaryotic RecQ-like helicases.....   | 91  |
| Figure 3-4  | Representation of protein domains predicted in TbMUS81.....   | 92  |
| Figure 3-5  | Multiple sequence alignment of MUS81 proteins.....  | 93  |
| Figure 3-6  | Neighbour joining tree of eukaryotic ERCC4-containing proteins ...  | 95  |
| Figure 3-7  | Representation of protein domains predicted in TbPIF6 .....   | 96  |
| Figure 3-8  | Multiple sequence alignment of Pif1 family helicases .....  | 96  |
| Figure 3-9  | Neighbour joining tree of eukaryotic Pif1 family helicases .....  | 99  |
| Figure 3-10 | Schematic of knockout constructs and targeting strategy .....   | 101 |
| Figure 3-11 | Diagnostic PCRs to confirm knockout of target genes .....   | 103 |
| Figure 3-12 | Confirmation by PCR of knockout of RECQ2, MUS81 and PIF6.....   | 104 |
| Figure 3-13 | RT-PCR confirmation of <i>recq2</i> , <i>mus81</i> and <i>pif6</i> mutants.....                             | 106 |
| Figure 3-14 | <i>In vitro</i> growth analysis of <i>recq2</i> , <i>mus81</i> and <i>pif6</i> mutants .....                | 108 |
| Figure 3-15 | Cell cycle analysis of <i>recq2</i> , <i>mus81</i> and <i>pif6</i> mutants.....                             | 109 |
| Figure 3-16 | <i>In vitro</i> growth of <i>recq2</i> and <i>pif6</i> mutants in the presence of DNA damaging agents ..... | 112 |
| Figure 3-17 | Clonal survival of <i>recq2</i> , <i>mus81</i> and <i>pif6</i> mutants in the presence of MMS .....         | 114 |
| Figure 3-18 | Clonal survival of <i>recq2</i> , <i>mus81</i> and <i>pif6</i> mutants in the presence of phleomycin .....  | 116 |
| Figure 3-19 | Clonal survival of <i>recq2</i> , <i>mus81</i> and <i>pif6</i> mutants in the presence of hydroxyurea ..... | 118 |
| Figure 3-20 | Schematic of re-expresser construct .....   | 120 |
| Figure 3-21 | PCR confirmation of MUS81 and PIF6 re-expresser cell lines.....   | 121 |
| Figure 3-22 | RT-PCR confirmation of MUS81 re-expresser cell lines.....   | 122 |
| Figure 3-23 | <i>In vitro</i> growth of MUS81 and PIF6 re-expresser cell lines.....                                       | 123 |
| Figure 3-24 | Clonal survival of re-expresser lines in the presence of DNA damaging agents .....                          | 124 |
| Figure 3-25 | Diagram of C-terminal 12myc tagging construct and strategy. ...   | 126 |
| Figure 3-26 | RECQ2 N-terminal 12myc tagging construct and strategy.....  | 128 |
| Figure 3-27 | Western blot of 12myc-tagged RECQ2, MUS81 and PIF6.....   | 130 |
| Figure 3-28 | Confirmation of 12myc/- cells by western blot and PCR.....  | 132 |

|             |   |     |
|-------------|---|-----|
| Figure 3-29 | Clonal survival of <i>myc</i> <sup>-/-</sup> cells in the presence of MMS .....   | 134 |
| Figure 3-30 | Representative examples of 12 <i>myc</i> RECQ2 localisation .....   | 136 |
| Figure 3-31 | Representative examples of MUS81 12 <i>myc</i> localisation .....   | 139 |
| Figure 3-32 | Representative examples of PIF6 12 <i>myc</i> localisation.....   | 141 |
| Figure 4-1  | I-SceI target sequence organisation in HR1 and HRES cell lines....  | 154 |
| Figure 4-2  | PCR confirmation of <i>RECQ2</i> knockout in HR1 and HRES cell lines..  | 157 |
| Figure 4-3  | PCR confirmation of <i>MUS81</i> knockout in HR1 and HRES cell lines..  | 158 |
| Figure 4-4  | Confirmation by PCR of <i>PIF6</i> knockout in HR1 and HRES cell lines.   | 159 |
| Figure 4-5  | Confirmation by RT-PCR of <i>recq2</i> and <i>mus81</i> knockout cells in HR1<br>and HRES cell line .....                           | 161 |
| Figure 4-6  | Repair of I-SceI-induced DSB in HR1 cells.....  | 163 |
| Figure 4-7  | Clonal survival following I-SceI induction in <i>HR1 recq2</i> mutants ...  | 165 |
| Figure 4-8  | Repeat of clonal survival assay of HR1 <i>recq2</i> mutants following I-SceI<br>DSB induction.....                                  | 167 |
| Figure 4-9  | Clonal survival following I-SceI induction in <i>HR1 mus81</i> mutants. .   | 169 |
| Figure 4-10 | Clonal survival following I-SceI induction in <i>HR1 pif6</i> mutants. ..   | 170 |
| Figure 4-11 | Primer binding sites for PCR analysis of HRES I-SceI assay clones.  | 172 |
| Figure 4-12 | Clonal survival and <i>ESAG1</i> and <i>VSG221</i> loss following I-SceI<br>induction in <i>HRES recq2</i> mutants.....             | 175 |
| Figure 4-13 | Repeat of analysis of clonal survival following I-SceI induction in<br><i>HRES recq2</i> mutants .....                              | 176 |
| Figure 4-14 | Clonal survival and <i>ESAG1</i> and <i>VSG221</i> loss following I-SceI<br>induction in <i>HRES mus81</i> mutants.....             | 179 |
| Figure 4-15 | Analysis of clonal survival and <i>ESAG1</i> and <i>VSG221</i> loss following<br>I-SceI induction in <i>HRES pif6</i> mutants ..... | 182 |
| Figure 4-16 | Attempt by PCR to analyse <i>VSG221</i> genomic location.....   | 183 |
| Figure 4-17 | Combined data of two repeats of clonal survival following I-SceI<br>induction in <i>HRES pif6</i> mutants.....                      | 184 |
| Figure 4-18 | DAPI cell cycle analysis of <i>HRES pif6</i> mutants .....  | 188 |
| Figure 4-19 | Cell cycle analysis of <i>HR1 pif6</i> mutants following I-SceI induction.  | 189 |
| Figure 4-20 | Analysis of <i>VSG221</i> expression in <i>HRES pif6</i> mutants following I-SceI<br>induction .....                                | 190 |
| Figure 4-21 | Analysis of $\gamma$ H2A expression in <i>HRES pif6</i> mutants.....  | 193 |
| Figure 4-22 | Analysis of $\gamma$ H2A expression in <i>HR1 pif6</i> mutants .....  | 197 |
| Figure 4-23 | Cell line <i>GFP221hygTK</i> used for VSG switching experiments .....   | 200 |
| Figure 4-24 | Strategy for determining VSG switching mechanisms.....  | 202 |
| Figure 4-25 | Constructs used in the generation of the <i>GFP221hygTK</i> VSG<br>switching cell line .....  | 205 |
| Figure 4-26 | PCR primer binding sites of PCR confirm <i>GFP221hygTK</i> cell line.   | 206 |
| Figure 4-27 | Confirmation of <i>GFP221hygTK</i> cell line by PCR and western blot.   | 206 |
| Figure 4-28 | PCR and western blot analysis of <i>GFP221hygTK recq2</i> mutants...  | 209 |
| Figure 4-29 | PCR and western blot analysis of <i>GFP221hygTK mus81</i> mutants..   | 210 |
| Figure 4-30 | PCR and western blot analysis of <i>GFP221hygTK pif6</i> mutants.....   | 212 |
| Figure 4-31 | Switching rate and profile in first <i>RECQ2</i> VSG switching assay....  | 216 |
| Figure 4-32 | PCR analysis of first <i>RECQ2</i> VSG switching experiment .....   | 217 |
| Figure 4-33 | Western analysis of first <i>RECQ2</i> VSG switching experiment .....   | 218 |
| Figure 4-34 | Switching rate and profile of <i>recq2</i> mutants in a second VSG<br>switching experiment .....                                    | 222 |
| Figure 4-35 | PCR analysis of second <i>RECQ2</i> VSG switching experiment .....  | 223 |
| Figure 4-36 | Western analysis of a second <i>RECQ2</i> VSG switching experiment..  | 225 |
| Figure 4-37 | Switching rate and switching profile of <i>recq2</i> mutants from a<br>combined dataset.....  | 230 |
| Figure 4-38 | Switching rate and switching profile of <i>pif6</i> mutants .....   | 233 |

|             |  |     |
|-------------|--|-----|
| Figure 4-39 | PCR analysis of <i>PIF6</i> VSG switching experiment .....   | 235 |
| Figure 4-40 | Western analysis of <i>PIF6</i> VSG switching experiment .....   | 236 |
| Figure 5-1  | Representation of TbRECQ1 protein sequence .....   | 252 |
| Figure 5-2  | Neighbour-joining tree of TbRECQ1 and other RecQ-like helicases .....                                      | 256 |
| Figure 5-3  | Schematic of <i>RECQ1</i> knockout constructs and knockout strategy..                                      | 257 |
| Figure 5-4  | PCR analysis of putative <i>recq1</i> <sup>-/-</sup> clones.....   | 260 |
| Figure 5-5  | Schematic of RNAi plasmid pGL2084.....   | 263 |
| Figure 5-6  | Schematic of RECQ1 RNAi construct, Gateway cloning and 2T1 transformation .....                            | 264 |
| Figure 5-7  | Confirmation of RECQ1 RNAi plasmid.....  | 266 |
| Figure 5-8  | PCR of <i>RECQ1</i> RNAi cassette integration in 2T1 cells .....   | 267 |
| Figure 5-9  | <i>In vitro</i> growth following <i>RECQ1</i> RNAi induction.....  | 268 |
| Figure 5-10 | Cell types observed 48 hours post- <i>RECQ1</i> RNAi induction .....                                       | 270 |
| Figure 5-11 | Cell cycle analysis following <i>RECQ1</i> RNAi induction .....  | 272 |
| Figure 5-12 | Analysis of <i>in vitro</i> growth in the presence of MMS following <i>RECQ1</i> RNAi induction.....       | 273 |
| Figure 5-13 | Immunofluorescent analysis of nuclear $\gamma$ H2A intensity following <i>RECQ1</i> RNAi induction.....    | 275 |
| Figure 5-14 | Schematic of the <i>RECQ1</i> 12myc plasmid.....   | 277 |
| Figure 5-15 | Western blot confirmation of RECQ1-12myc expression in RECQ1 12myc RNAi clones.....                        | 279 |
| Figure 5-16 | Confirmation of <i>RECQ1</i> knockdown by western blot analysis .....                                      | 279 |
| Figure 5-17 | Confirmation of RECQ1 12myc expression by western blot.....  | 281 |
| Figure 5-18 | Confirmation of successful <i>RECQ1</i> 12myc/ $\Delta$ <i>RECQ1</i> by PCR and western blot analysis..... | 282 |
| Figure 5-19 | Representative examples of attempts at RECQ1 12myc immunolocalisation.....                                 | 283 |
| Figure 7-1  | Individual <i>pif6</i> mutant MMS clonal survival assays.....  | 311 |
| Figure 7-2  | Diagnostic PCRs from first <i>RECQ2</i> HRES I-SceI assay .....  | 313 |
| Figure 7-3  | Diagnostic PCRs of survivors from <i>MUS81</i> HRES I-SceI assay .....                                     | 317 |
| Figure 7-4  | Diagnostic PCRs of survivors from first <i>PIF6</i> HRES I-SceI assay .....                                | 320 |

## Acknowledgements

Firstly, I would like to thank my supervisor Richard McCulloch for his guidance and relentless enthusiasm throughout this work, as well as his always open door and subsequent patience with my endless questions.

Thank you to all in the McCulloch lab and colleagues at the WTCMP past and present who made working at the WTCMP a pleasure. I will miss teatime and the often unexpected turns that conversations took. Special thanks go to Catarina, for tissue culture Disney marathons, afternoon coffee runs, generally being my partner in crime and helping me preserve my sanity in my last few months of lab work. Thank you to Marko, for his invaluable scientific advice throughout my PhD and for his many extraordinary and entertaining stories. Also to Craig, for much appreciated technical help, especially with VSG switching assays, as well as putting up with the aforementioned Disney marathons. Thank you to Alex - I've yet to find an administrative problem she can't fix.

Outwith Glasgow, advice on I-Scel cell lines and reagents from Prof. David Horn and Dr Lucy Glover at the University of Dundee were invaluable in producing this work. Thanks also to Dr Gloria Rudenko and her lab at Imperial College London for the plasmids to construct VSG switching lines and guidance on assays.

Thank you to the Directors of the Wellcome Trust 4-year PhD programme, Prof. Darren Monckton and Dr Olwyn Byron, for the opportunity to pursue this work and the Wellcome Trust for funding.

I'd also like to thank my family and friends for their encouragement and positivity, even when they had no clue what I was talking about. Finally, I would like to thank Steph for her support, unceasing optimism throughout the writing of this thesis, never failing to provide distraction and cabin fever relief when needed, even when it meant going for walks in -30°C weather and lastly, for much needed (and appreciated) proofreading.

“Tea, Earl Grey, hot!”

Captain Jean-Luc Picard

## **Author's declaration**

I declare that this thesis and the results presented within are entirely my own work except where otherwise stated. No part of this thesis has been previously submitted for a degree at any university.

Rebecca Devlin

## Definitions and Abbreviations

|                   |   |
|-------------------|---|
| ATP               | Adenosine triphosphate                          |
| BES               | Blood stream form expression site               |
| BIR               | Break-induced replication                       |
| BLAST             | Basic Local Alignment Search Tool               |
| BLE               | Bleomycin                                       |
| BLM               | Bloom syndrome protein                          |
| bp                | Base pairs                                      |
| BSD               | Blasticidin                                     |
| BSF               | Blood stream form                               |
| BVSG              | Blood stream form VSG                           |
| cDNA              | Complementary DNA                               |
| CFS               | Common fragile site                             |
| CNS               | Central nervous system                          |
| CTR               | Co-transposed region                            |
| DAPI              | 4',6-diamidino-2-phenylindole                   |
| del               | Deletion  |
| DGC               | Directional gene cluster                        |
| dH <sub>2</sub> O | Distilled water                                 |
| dHJ               | Double Holliday Junction                        |
| DIC               | Differential interference                       |
| DNA               | Deoxyribonucleic acid                           |
| DNA-PKcs          | DNA dependent protein kinase, catalytic subunit |
| Dnase             | Deoxyribonuclease                               |
| dNTP              | Deoxynucleotide triphosphate                    |
| DSB               | Double strand break                             |
| DSBR              | DSB repair                                      |
| dsDNA             | Double stranded DNA                             |
| dsRNA             | Double stranded RNA                             |
| EDTA              | Ethylenediaminetetraacetic acid                 |
| EF1 $\alpha$      | Elongation factor 1-alpha                       |
| eGFP              | Enhanced green fluorescent protein              |
| ES                | Expression site                                 |
| ESAG              | Expression site associated genes                |



|       |   |
|-------|---|
| FACS  | Fluorescence activated cell sorting                   |
| FBS   | Foetal bovine serum                                   |
| FITC  | Fluorescein isothiocyanate                            |
| G4    | Guanine quadruplex                                    |
| GC    | Gene conversion                                       |
| gDNA  | Genomic DNA   |
| GFP   | Green fluorescent protein                             |
| GPI   | Glycophosphatidylinositol                             |
| HA    | Human influenza hemagglutinin                         |
| HAT   | Human African trypanosomiasis                         |
| HJ    | Holliday junction                                     |
| HMI   | Hirumi's modified Isocove's medium                    |
| HR    | Homologous recombination                              |
| HRDC  | Helicase and RNaseD C-terminal                        |
| HU    | Hydroxyurea   |
| HYG   | Hygromycin  |
| IMDM  | Isocove's modified Dulbecco's medium                  |
| kb    | Kilobase pairs  |
| kda   | Kilo-dalton   |
| kDNA  | Kinetoplast DNA                                       |
| LB    | Luria Bertani   |
| MACS  | Magnetic activated cell sorting                       |
| MMEJ  | Microhomology-mediated end joining                    |
| MMS   | methyl methanesulfonate                               |
| MOPS  | 3-(N-morpholino) propanesulfonic acid                 |
| MRN   | Mre11-Rad50-Nbs1                                      |
| mRNA  | Messenger RNA   |
| MRX   | Mre11-Rad50-Xrs2                                      |
| MVSG  | Metacyclic VSG  |
| MES   | Metacyclic form expression site                       |
| NADH  | Nicotinamide adenine dinucleotide (reduced)           |
| NADPH | Nicotinamide-adenine dinucleotide phosphate (reduced) |
| NEO   | Neomycin  |
| NHEJ  | Non-homologous end joining                            |
| NTP   | Nucleotide triphosphate                               |

|           |   |
|-----------|---|
| ORF       | Open reading frame                              |
| PAGE      | Polyacrylamide gel electrophoresis              |
| PBS       | Phosphate buffered saline                       |
| PBST      | PBS-tween                                       |
| PCF       | Procyclic form                                  |
| PCR       | Polymerase chain reaction                       |
| PDB       | Protein Data Bank                               |
| PFA       | Paraformaldehyde                                |
| PUR       | Puromycin                                       |
| RFP       | Red fluorescent protein                         |
| RIF       | RNA interference factor                         |
| RISC      | RNA-induced silencing complex                   |
| RNA       | Ribonucleic acid                                |
| RNA pol I | RNA polymerase I                                |
| RNAi      | RNA interference                                |
| RNAPI     | RNA polymerase I                                |
| RPA       | Replication protein A                           |
| RQC       | RecQ C-terminal                                 |
| RRNA      | Ribosomal RNA                                   |
| RT        | Reverse transcriptase                           |
| RT-PCR    | Reverse transcriptase polymerase chain reaction |
| RTR       | RecQ-Top3-Rmi1                                  |
| RXP       | Re-expresser                                    |
| SAP       | SAF-A/B, Acinus and PIAS                        |
| SDS       | Sodium dodecyl sulphate                         |
| SDSA      | Synthesis-dependent strand annealing            |
| SEM       | Standard error of the mean                      |
| SF        | Superfamily                                     |
| siRNA     | Small interfering RNA                           |
| SMARD     | Single molecule analysis of replicated DNA      |
| SOC       | Super optimal broth supplemented with glucose   |
| SSA       | Single strand annealing                         |
| SSB       | Single strand break                             |
| ssDNA     | Single stranded DNA                             |
| TAE       | Tris acetic acid EDTA buffer                    |

|      |                              |
|------|------------------------------|
| Tb   | Trypanosoma brucei           |
| Tet  | Tetracycline                 |
| TetR | Tetracycline repressor       |
| TK   | Thymidine kinase             |
| UTR  | Untranslated region          |
| UV   | Ultraviolet                  |
| VSG  | Variant surface glycoprotein |
| WRN  | Werner syndrome              |
| WT   | Wild type                    |
| XO   | Crossover                    |

# **Chapter 1**

## **Introduction**

# 1 Introduction

## 1.1 General Introduction

*Trypanosoma brucei* is a eukaryotic parasite within the *Trypanosomatidae* family, a group of protozoan parasites belonging to the *Kinetoplastida* order. Kinetoplastids are distinguished by possession of a kinetoplast, a concatenated disc-shaped network of circular DNA molecules (kDNA) that comprise the genome of the single mitochondrion found in each cell, located adjacent to the flagellar basal body (Lukes *et al.*, 2002). Kinetoplastid organisms possess a number of unusual biological processes compared to other eukaryotes including extensive RNA-editing, a process that involves nucleotide insertion and deletion of mitochondrial (kDNA) transcripts. This is necessary in order to produce mature transcripts for translation. The *Trypanosomatidae* family includes the *Leishmania* and *Trypanosoma* genera, which contain a number of parasitic species that infect humans as well as other vertebrates (Simpson *et al.*, 2006). *T. brucei* causes human African trypanosomiasis (HAT), commonly known as African sleeping sickness, in humans and nagana in cattle. *T. brucei* is transmitted by the tsetse fly vector, thus restricting its distribution to areas of Africa the tsetse fly inhabits.

African trypanosomes live exclusively extracellularly in the bloodstream and tissue fluids of mammals. This extracellular lifestyle contrasts with the intracellular *Trypanosoma cruzi*, a trypanosome species found in South America causing Chagas disease. The extracellular lifestyle of African trypanosomes is achieved by antigenic variation, the periodic switching in expression of their variant surface glycoprotein (VSG) coat, which enables the parasite population to continually evade destruction by the host's immune system (discussed in Section 1.2). This leads to chronic, and if untreated fatal, *T. brucei* infections. These *Trypanosoma* species are of substantial health and economic importance due to their impact upon humans, directly and indirectly through loss of livestock. *T. brucei* is the African trypanosome species on which the greatest amount of research has been conducted, facilitated by its amenability to laboratory culture.

The *T. brucei* species is divided into the subspecies *T. brucei brucei*, *T. brucei gambiense* and *T. brucei rhodesiense*. These subspecies also have a different geographical distribution and infection profile: *T. brucei rhodesiense* is present primarily in southern and eastern Africa and causes a fast onset HAT, while *T. brucei gambiense* is present in western and central Africa and causes a slow onset, chronic form of HAT (Barrett *et al.*, 2003). In contrast, *T. brucei brucei* is not infective to humans due to parasite uptake of trypanosome lytic factor (TLF), a toxic high-density lipoprotein (HDL) from human serum (Hajduk *et al.*, 1992; Harrington *et al.*, 2009), composed of apolipoprotein A-I (apo A-I) and haptoglobin related protein (Hpr) (Shiflett *et al.*, 2005). *T. brucei brucei* is closely related to the human-infective subspecies of *T. brucei* and shares many important features with them, including VSG switching. It is therefore frequently used as a laboratory model for human trypanosomiasis. *T. brucei brucei* is however infective to mice, thus providing an *in vivo* model for trypanosomiasis studies.

### 1.1.1 *Trypanosoma brucei* life cycle

For completion of its life cycle (Fig. 1-1), *T. brucei* must survive and proliferate in both mammals and the tsetse flies (genus *Glossina*). The disparate environments of the mammalian bloodstream and the tsetse fly has necessitated *T. brucei* to evolve a complex life cycle with distinct stages adapted to the environments it inhabits.

Mammalian hosts are infected via tsetse fly bites. The saliva of infected tsetse flies contains metacyclic form trypanosomes, which are transmitted to the mammal when the fly takes its blood meal. Metacyclic form trypanosomes are pre-adapted to infect mammals and cell cycle arrested until they enter the mammalian bloodstream (Matthews & Gull, 1997; Shapiro *et al.*, 1984).

Metacyclic cells possess a metacyclic variant surface glycoprotein (MVSG) coat to protect them from the immune response of the mammalian host they enter (Tetley *et al.*, 1987; Turner *et al.*, 1988).

In the mammalian bloodstream, a series of morphological and biochemical changes in the metacyclic form cells occurs, leading to differentiation into 'long slender' bloodstream form (BSF) cells adapted for survival and proliferation in

the mammalian bloodstream. The MVSG surface coat is replaced with a BSF VSG (BVSG) coat to protect them from the innate and adaptive immune responses. Cells proliferate rapidly in the bloodstream. The characteristic feature of BSF cells is the periodic switching of the single VSG type on the surface of a cell to an antigenically distinct VSG - the basis for immune evasion by antigenic variation that is shared by many pathogens (discussed in Section 1.2).

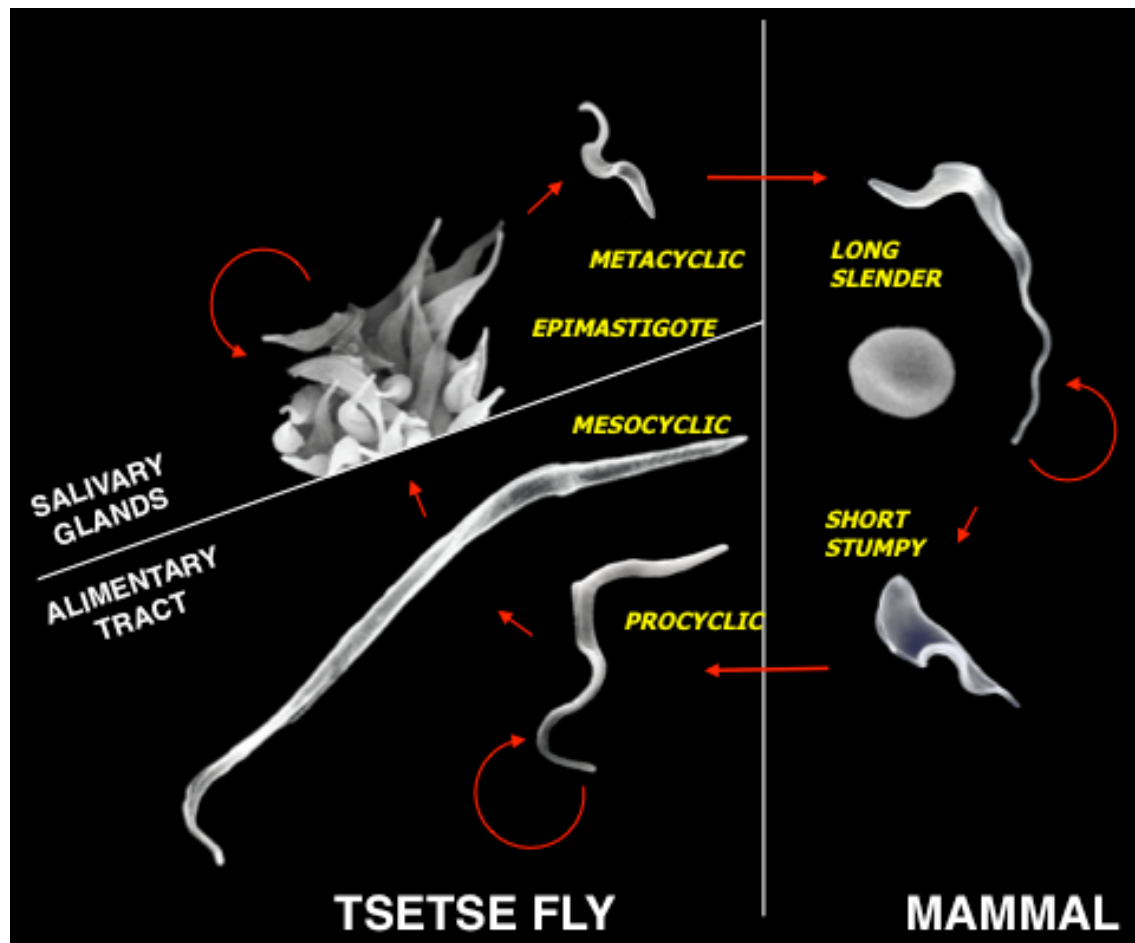


Figure 1-1 *T. brucei* life cycle.

The life cycle stages of *T. brucei* are shown, in the form of scanning electron micrograph (SEM) images (to scale). The host animal of each stage and location of life cycle stages within the tsetse fly are indicated. The name of each life cycle stage is shown, with straight arrows indicating progression from one stage to another and curved arrows indicating proliferative stages. Figure from Barry & McCulloch (2001).

Long slender BSF parasites differentiate into 'short stumpy' BSF trypanosomes, reviewed in Rico *et al.* (2013). As parasitaemia increases, long slender form cells differentiate into short stumpy form cells that are cell cycle arrested (Reuner *et al.*, 1997; Vassella *et al.*, 1997). Cell cycle arrested short stumpy BSF cells die if they are not transmitted to the insect vector (Turner *et al.*, 1995), which occurs by ingestion when the fly takes a blood meal. Trypanosomes pass

into the midgut of the fly, where they differentiate into replicative procyclic form (PCF) cells, then into proventricular mesocyclic forms that migrate to the salivary glands. In the salivary glands epimastigotes and prometacyclics form, which differentiate into infective metacyclics. The salivary glands are also the location of trypanosome meiosis (Peacock *et al.*, 2014; Peacock *et al.*, 2011). Metacyclic cells are transmitted back to a mammal when the fly takes its blood meal. Though the mechanism driving the differentiation of PCF cells to metacyclic cells in the tsetse fly is unknown, it has been replicated *in vitro* by over expression of the RNA binding protein RBP6 (Kolev *et al.*, 2012), suggesting a role for RBPs in this process. PCF and BSF cells are the two *T. brucei* life cycle stages amenable to *in vitro* culture.

### **1.1.2 Human African Trypanosomiasis: prevalence, symptoms and treatment**

African sleeping sickness occurs in 36 sub-Saharan African countries, areas where the tsetse fly vector is present to facilitate transmission. Recent control efforts have resulted in fewer than 8000 reported cases in 2012, down from an estimated 300,000 cases in 1998 (World Health Organisation). Approximately 95% of HAT cases are due to *T. brucei gambiense*, with only a small minority caused by *T. brucei rhodesiense* (World Health Organisation).

Human African Trypanosomiasis (HAT) is classed as either stage 1 (haemolympathic) or stage 2 (cerebral), depending on whether the trypanosome parasites have entered the central nervous system (CNS). In Stage 1, trypanosomes proliferate in the blood and lymphatic tissue; it can last for weeks or years and can include asymptomatic periods (Jamonneau *et al.*, 2012; Songa *et al.*, 1991). Stage 2 HAT is characterised by trypanosome invasion of internal organs, including the CNS. This progression can take months or years in *T. brucei gambiense* infections but can occur within weeks in *T. brucei rhodesiense* infections (Barrett *et al.*, 2003).

There are currently only four drugs licensed for the treatment of HAT: pentamidine, suramin, melarsoprol and eflornithine (World Health Organisation). These drugs are associated with significant toxicity to the patient and/or dosing that is problematic for treatment of patients in under developed countries



(Milord *et al.*, 1992; Pepin & Milord, 1991; Wilkinson & Kelly, 2009). Additionally, the dense, antigenically variant VSG surface coat means development of a vaccine is not promising (La Greca & Magez, 2011).

### 1.1.3 *T. brucei* genome

The haploid genome of *T. brucei* (strain TREU927; GeneDB version 2013-08-28) is ~35 Mb in size and primarily comprised of 11 mega base-sized chromosomes. Genome sequencing (strain TREU927; GeneDB version 2013-08-28) revealed 12094 predicted protein-coding genes, including 1458 pseudogenes. The gene repertoire of strain Lister 427 is less than that of TREU 927, due to a greater degree of analysis having been performed on strain TREU927. Analysis of strain Lister 427, which has revealed 9302 predicted protein coding genes, including 524 pseudogenes and a total genome size of ~26 Mb (GeneDB, version 2010-10-20).

In their analysis of VSGs in the main contigs of the megabase chromosomes of TREU927, Marcello & Barry (2007) analysed 904 VSGs, whilst Cross *et al.* (2014) identified 2563 distinct VSG genes in the *T. brucei* Lister 427 genome. The 11 mega base chromosomes range in size from 0.9 Mb to >6 Mb and are numbered 1 to 11, 1 being the smallest and 11 the largest in size. However, chromosome size varies between and within strains, with differences of up to four-fold between chromosome homologue pairs in a single isolate (El-Sayed *et al.*, 2000). Approximately 20% of the genome encodes subtelomeric sequences, the majority of which are *T. brucei* specific, and a large number of these are involved in antigenic variation (Berriman *et al.*, 2005), with many of the telomeres linked to VSG bloodstream form expression sites (VSG BESs) (El-Sayed *et al.*, 2000).

Telomeres are nucleoprotein structures, located at the ends of chromosomes to protect them from processes such as exonucleolytic degradation and end-to-end fusions (Stewart *et al.*, 2012). Eukaryotic telomeres are usually composed of TG-rich simple repetitive sequences that are replicated by the specialised telomerase enzyme (Nandakurnar & Cech, 2013; Stewart *et al.*, 2012). Between the telomere and the core chromosome sequence lies a region known as the subtelomere. There is no absolute definition of what a subtelomere is, but it is agreed that subtelomere sequence is more complex than the simple repeats of

the telomere but lacks the degree of conservation between chromosomes that the chromosome core possesses, are highly polymorphic and contain an array of types of repeat element (Churikov, 2001; Louis & Vershinin, 2005). Unlike telomeres, which do not contain genes, subtelomeres encode genes with important cellular functions and human subtelomere instability is associated with disease (Sacconi *et al.*, 2013; Mefford *et al.*, 2012). A number of pathogenic organisms that undergo antigenic variation, including *T. brucei*, *Plasmodium falciparum* and *Borrelia burgdorferi*, express the genes encoding variant surface antigens from subtelomeric regions (Horn, 2014; Claessens *et al.*, 2014; Zhang *et al.*, 1997). Additionally, human subtelomeres encode human olfactory receptor genes (Mefford, 2001; Trask, 1998) and immunoglobulin heavy chains (Das *et al.*, 2008), further evidence for subtelomeres as important locations for diverse gene families.

Expression of the *T. brucei* genome is also unusual among eukaryotes. *T. brucei* is the only known eukaryotic organism in which RNA polymerase I (RNAPI) transcribes not only *rRNA* genes but also the protein-coding VSG and procyclin genes (Gunzl *et al.*, 2003). The gene organisation of trypanosomes is also unusual, with virtually all genes organised into large directional gene clusters (DGCs) containing tens of genes polycistronically transcribed by RNAPII (Imboden *et al.*, 1987). Polytranscriptional units are bounded by variant histones (Siegel *et al.*, 2009), transcriptional start sites are marked by histone H3K4 trimethylation (Wright *et al.*, 2010) and base J ( $\beta$ -d-glucosyl-hydroxymethyluracil, a modification of thymine) maps to transcriptional terminators (Reynolds *et al.*, 2014). Furthermore, the *T. brucei* genome displays a segregation of gene types, with house-keeping genes clustered in DGCs in the core of the megabase chromosomes, VSG gene sequences clustered in the subtelomeres and BESs proximal to the telomeres (Barry *et al.*, 2005). The association of genes involved in antigenic variation with telomeres, an organisation found in other organisms, has been suggested to be due to the ease with which they ectopically recombine, thereby promoting VSG recombination (Horn & Barry, 2005).

*T. brucei* also contain several intermediate chromosomes (200-900 Kb) and ~100 mini-chromosomes (50-150 kb) (El-Sayed *et al.*, 2000). The only genes found so far on mini-chromosomes encode VSG and expression site associated genes

(ESAGs), though they do not contain VSG BESs (Wickstead *et al.*, 2004). This suggests that for expression of VSGs located on mini-chromosomes, recombination of the gene into a VSG BES elsewhere is necessary. In contrast, many VSGs found in intermediate chromosomes are present in VSG BESs, some of which at least can be actively transcribed from VSG BESs (El-Sayed *et al.*, 2000).

## 1.2 *T. brucei* antigenic variation

Though *T. brucei* is not the only pathogenic organism to undergo antigenic variation, it is arguably has the most complex and extensive antigenic variation system. Antigenic variation is a common strategy for pathogenic organisms to prolong their survival in their host and enhance transmission and fitness (Barry & McCulloch, 2001; Deitsch *et al.*, 2009; Morrison *et al.*, 2009). Antigenic variation is a process exclusively concerned with evasion of host acquired immunity. In this process, an organism expresses a single surface antigen at one time and then switches among a series of antigenically distinct variants, each singly expressed over time. The formation of such ‘families’ of surface antigen genes can be achieved by a number of mechanisms. Recombination is a common mechanism of antigenic variation and is exclusively used in some pathogens such as *Anaplasma marginale* and *Borrelia burgdorferi*, which undergo antigenic variation of *Msp2* (major surface protein 2) and *VlsE*, a surface lipoprotein respectively, both reviewed in Palmer *et al.* (2009). *Neisseria* too uses gene conversion recombination (discussed in Section 1.3.2.1) to transfer transcriptionally silent *PilS* sequences to the single *PilE* expression site. This creates antigenic diversity in the gene encoding pilin, the major subunit of the *Neisseria* type IV pilus, reviewed in Cahoon & Seifert (2011). However, antigenic variation does not have to occur by recombination. Several pathogens utilise transcriptional control to switch antigen variants. *P. falciparum* switches between ~60 *var* genes, which encode erythrocyte membrane protein 1 (PfEMP1) proteins, by transcriptional control, reviewed in Guizetti & Scherf (2013). *Giardia lamblia* too, uses transcriptional control to switch between expression of ~200 variant surface proteins (VSPs), reviewed in Prucca *et al.* (2011).

Like *Neisseria*, *T. brucei* employs an antigenic variation system that uses multiple copies of a gene, the VSG, and switches between them. However, the *T. brucei* system is far more elaborate. While *Neisseria* possesses six variants of

the pilin gene (Hill & Davies, 2009), the *T. brucei* distinct VSG genes genome contains 2563 distinct VSG genes (Cross *et al.*, 2014) that are expressed on the *T. brucei* cell surface. In addition, *T. brucei* possesses >1 site of VSG expression, the VSG BESs, meaning that VSG switching is not limited to recombination, but can also employ transcriptional control mechanisms. This huge VSG archive is much larger than any antigenically variant gene archive described elsewhere in nature (Morrison *et al.*, 2009).

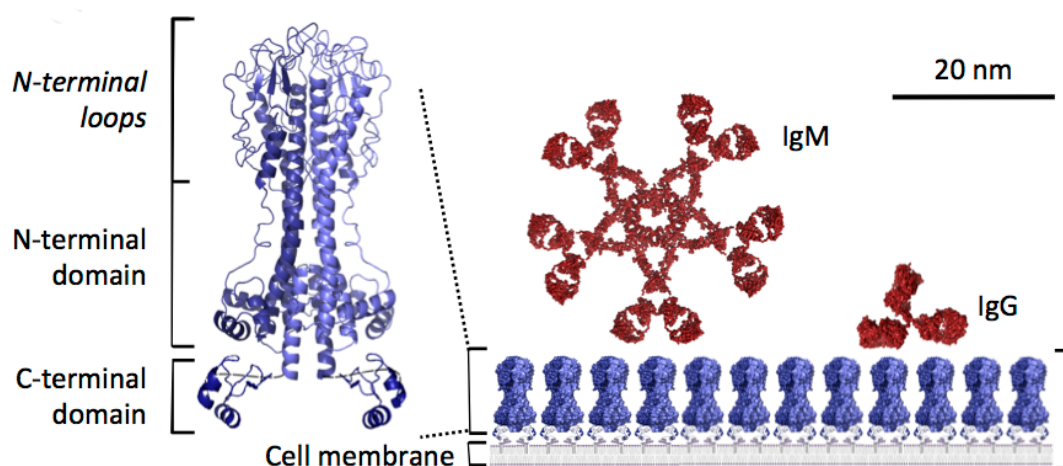
It has until recently been thought that the VSG is the only antigen accessible to the adaptive immune system due to the dense packing of the VSGs in the coat (Fig. 1-6), therefore concealing the invariant surface molecules beneath as well as preventing complement-mediated lysis machinery from accessing the cell membrane. The host can thus only generate an antibody response to the VSG expressed by the majority of the parasite population. However, a small subset of cells switch expression to an antigenically different VSG and this subpopulation escapes the immune response targeted to the parental population. This pattern is repeated, and together with the density-dependent differentiation of cells from long slender forms to short stumpy forms, produces characteristic parasitaemic waves and a prolonged infection (MacGregor *et al.*, 2011; MacGregor *et al.*, 2012). The view that the VSG coat shields all invariant *T. brucei* antigens has been challenged recently however, by the discovery that a *Trypanosoma congolense* haptoglobin-hemoglobin receptor is more elongated than a VSG protein and thus may extend beyond the VSG coat (Higgins *et al.*, 2013).

### **1.2.1 *T. brucei* variant surface glycoproteins (VSGs)**

Homodimers of variant surface glycoprotein (VSG) cover the cell surface of *T. brucei*, forming a dense monolayer and hiding the invariant surface molecules beneath from the host immune system (Fig. 1-2). VSGs are 400-500 amino acids in length and the majority are between 420 and 460 amino acids long (Hutchinson *et al.*, 2003). The C-terminal region is buried in the cell surface where it mediates attachment to the plasma membrane via a GPI anchor (Ferguson *et al.*, 1988), while the N-terminal domain points outward, exposed to the immune system. Although VSGs have a conserved structure dominated by alpha helices, they typically share less than 25% sequence identity with one

another (Hutchinson *et al.*, 2003). Variation is particularly high in the hyper variable N-terminal region of 350-400 residues that are exposed to the immune system (Berriman *et al.*, 2005; Hutchinson *et al.*, 2003).

*T. brucei* contains an estimated 1600 VSGs in its megabase chromosomes, with most of these genes located in subtelomeric arrays (Berriman *et al.*, 2005), as well as a further ~200 on the mini-chromosomes and ~20 VSGs located in BESs (see below) (Berriman *et al.*, 2005; Marcello & Barry, 2007; Wickstead *et al.*, 2004). Analysis of 940 silent VSG genes in TRUE927 by Marcello and Barry (2007) revealed that the vast majority of these were not intact genes capable of being expressed. Fewer than 5% were fully intact VSGs and the remainder were full length or partial VSG pseudogenes. Analysis of the VSGs of Lister 427 (Cross *et al.*, 2014) revealed a similar picture: 80% of the 2563 distinct complete and partial VSGs analysed were incomplete or pseudogenes. The large pseudogene repertoire acts as a resource for the generation of novel full-length mosaic VSGs through recombination reactions.



**Figure 1-2 The VSG surface coat of *T. brucei***

The VSG coat covers the surface of the trypanosome, physically hindering immune access to invariant antigens beneath. The VSG coat is composed of VSG dimers, (blue) composed of a C-terminal domain, N-terminal domain and N-terminal loops. The N-terminal loops are exposed to the immune system and are the most variable regions of the VSG. Figure adapted from James Hall, PhD Thesis, 2012 and originally assembled using PDB structures 1vsg, 1xu6, 1rcj and 1igt, visualized using Pymol (Schrödinger, LLC).

Trypanosomes also encode up to 27 distinct metacyclic variant surface glycoproteins (MVSGs) that are expressed on the surface of metacyclic cells (Turner *et al.*, 1988) in the tsetse fly salivary gland. This MVSG remains transcriptionally active until after the metacyclic from cell passes into a

mammalian host and begins to differentiate into the BSF (Ginger *et al.*, 2002). The MVSG coat is then lost and replaced by a bloodstream VSG (BVSG) coat.

### 1.2.2 VSG expression sites

VSGs are transcribed from expression sites (ESs), polycistronic loci providing a mechanism for monoallelic VSG expression. VSGs can only be expressed if they are recombined into the ES and only one ES is actively transcribed at any one time, resulting in expression of a single VSG in one cell. Trypanosomes possess two types of ES: VSG blood stream expression sites (VSG BESs) and metacyclic expression sites (MESs). *T. brucei brucei* EATRO 2340 possesses 23 distinct VSG BESs (Young *et al.*, 2008), while Hertz-Fowler *et al.* (2008) identified 15 VSG BESs in the Lister 427 strain. One suggestion for why *T. brucei* possesses multiple VSG BESs is that they may serve as sites for the construction of novel mosaic VSGs through recombination reactions (Marcello & Barry, 2007). Another possibility is that multiple VSG BESs allow for adaptation to growth in different hosts through expression of different expression site associated genes (ESAGs) (Bitter *et al.*, 1998; Cordon-Obras *et al.*, 2013; Pays *et al.*, 2001). VSG BESs (Fig. 1-3) are between 40 kb and 70 kb in length (Becker *et al.*, 2004) and are comprised of a telomere-proximal VSG, invariably flanked upstream by a number of 70 bp repeats as well as a number of promoter-proximal expression site associated genes (ESAGs) and a region of repeats known as 50 bp repeats. The 70 bp repeat region varies in size between 0.1 kbp and 7.0 kbp (McCulloch & Horn, 2009) and is not limited to the VSG BES, as over 90% of subtelomeric array VSGs have at least one 70 bp repeat 1-2 kbp upstream (Marcello & Barry, 2007). These repeats form the boundaries of gene conversion events that recombine intact, silent VSGs into active VSG BESs for expression (Alsford *et al.*, 2009). Thus, these repeats provide 5' sequence similarity between highly dissimilar VSGs, though it has been suggested that they may provide a more active function in recombination reactions (discussed in Section 1.2.4).

ESAGs are co-transcribed with VSGs from the VSG BES; trans-splicing and polyadenylation is used to process the primary VSG BES transcript to produce monocistronic mRNAs. The type of ESAGs present in a VSG BES varies, as does the number, though the relative order of the genes is rather conserved (Hertz-Fowler *et al.*, 2008; McCulloch & Horn, 2009). Relatively little is known about

the function of *ESAGs*, although a number of functions have been proposed and it has been suggested that they may serve to allow for adaptation to different host environments (Bitter *et al.*, 1998; Cordon-Obras *et al.*, 2013; Pays *et al.*, 2001).



**Figure 1-3 Generalised schematic of a bloodstream VSG BES**  
A typical bloodstream VSG expression site (BES) is shown. The 50 bp repeat region (hatched box) is shown, followed by the promoter is shown (flag), a number of expression site associated genes (*ESAGs*) (turquoise boxes), the 70 bp repeat region (hatched box) and the VSG located most telomerically (pink box). Telomeric repeats are indicated by triangles. Not to scale.

Transcription of the VSG BES (and MES) is carried out by RNAPI, which appears to initiate transcription at both active and inactive VSG BES loci, though transcription is only fully processive at the single active VSG BES (Vanhamme *et al.*, 2000), leading to monoallelic expression of only one VSG and its associated *ESAGs*. In addition, there is evidence that there is also VSG BES regulation at the level of transcription initiation (Nguyen *et al.*, 2014) and that VSG BESs are also regulated epigenetically (Figueiredo & Cross, 2010; Lopez-Farfan *et al.*, 2014; Stanne & Rudenko, 2010). The expression site body (ESB) is another important explanation for monoallelic VSG expression. The ESB is an extranucleolar RNA polymerase I-containing body that is associated specifically with the active VSG BES (Navarro & Gull, 2001). This association of the ESB with a specific VSG BES is heritable and crucial for the inheritance of monoallelic VSG expression. The means by which trypanosomes achieve monoallelic VSG expression has been recently reviewed by Günzl *et al.* (2015) and Glover *et al.* (2013b). *T. brucei* also contains monoallelically expressed metacyclic expression sites (MES). Although MESs are also transcribed by RNAPI, their promoter sequences are different to VSG BESs and they do not contain functional *ESAGs* (Ginger *et al.*, 2002).

### 1.2.3 Mechanisms of VSG switching in *T. brucei*

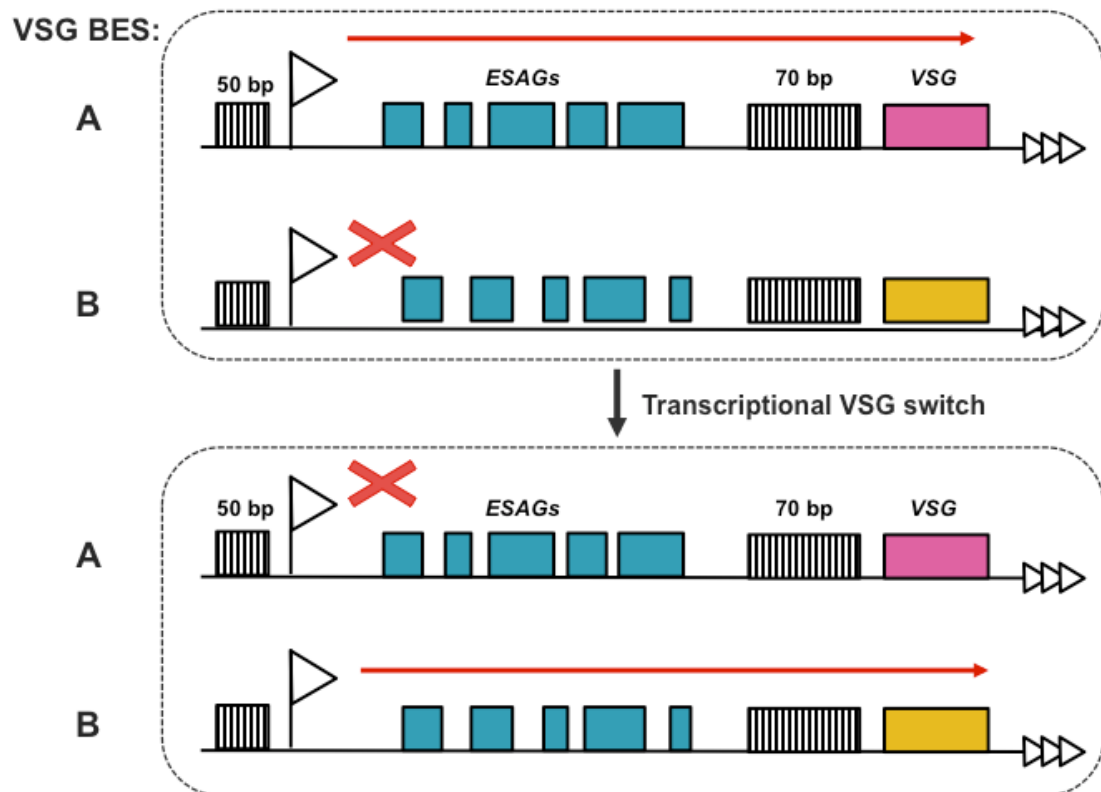
*T. brucei* possesses multiple mechanisms to facilitate the switching of expression from one VSG to another. These are categorised as either transcription-based switching or recombination-based switching.

### 1.2.3.1 Transcriptional VSG switching

Transcriptional VSG switching, also termed *in situ* switching (Fig 1-4), involves the cessation of transcription from the single active BES accompanied by the activation of another single VSG BES, harbouring a distinct VSG, and resulting in a change in VSG expression. Attempts to isolate *T. brucei* cells with two activate VSG BESs showed that this is a highly unstable state (Chaves *et al.*, 1999), illustrating the rigorous control of this process. Transcriptional control thus appears to be coordinated such that the switching ‘off’ and ‘on’ of VSG BESs are events dependent on one another (Chaves *et al.*, 1999).

Work by Figueiredo *et al.* (2008) showed that epigenetic mechanisms are involved in transcriptional VSG switching. Figueiredo *et al.* showed that *DOT1B*, a histone methyltransferase that methylates H3K79, is involved in transcriptional switching. Deletion of *DOT1B* resulted in de-repression of silent VSGs at telomeric loci, indicating that histone methylation is important in maintaining monoallelic VSG transcription. Additionally, nucleosomes are depleted at the active VSG BES compared to inactive VSG BESs (Figueiredo & Cross, 2010), indicating a role for chromatin remodelling in VSG BES regulation. Depletion of cohesin results in loss of monoallelic VSG expression and also causes an increase in the rate of VSG transcriptional switching (Landeira *et al.*, 2009), suggesting a possible link to DNA replication or cell division. It is also possible that DNA repair plays a role in transcriptional switching since induction of DNA damage results in the transcriptional activation of silent VSG BESs (Shedder *et al.*, 2004). Utilisation of the much larger number of VSGs outside the VSG BES, by recombinational VSG switching, is required for the chronic and fatal infections caused by *T. brucei*.





**Figure 1-4 Transcriptional VSG switching.**  
Two VSG bloodstream expression sites (VSG BESs) are shown (A & B). Initially, the VSG in BES (A) (pink box) is expressed, indicated by the red arrow and the VSG in BES (B) is inactive, indicated by a red cross. The VSG in BES (B) (orange box) can be activated by turning off transcription of VSG BES (A) and turning on transcription of VSG BES (B). Not to scale.

### 1.2.3.2 Recombinational VSG switching

VSG switching via recombination is the most common form of VSG switching, at least in pleomorphic lines, whereas monomorphic lines are believed to have a bias towards transcriptional switching (Aitchison *et al.*, 2005; Robinson *et al.*, 1999) - see Section 1.3.2 for discussion of homologous recombination.

Pleomorphic *T. brucei* cell lines are those that undergo differentiation from long slender BSF cells to short stumpy BSF cells. In contrast, monomorphic lines do not readily undergo this differentiation. It has been suggested that recombination is repressed in the monomorphic lines used in the laboratory such as Lister 427, though this is debated (Barry, 1997). Recombination reactions allow the movement of silent VSGs from elsewhere in the genome into active VSG BES, leading to transcription of the new VSG. Silent VSGs in the subtelomeric tandem arrays in the megabase chromosomes are the main source of VSGs moving into the active VSG BES. However, silent VSGs in the mini-chromosome subtelomeres and in silent VSG BESs in the megabase and

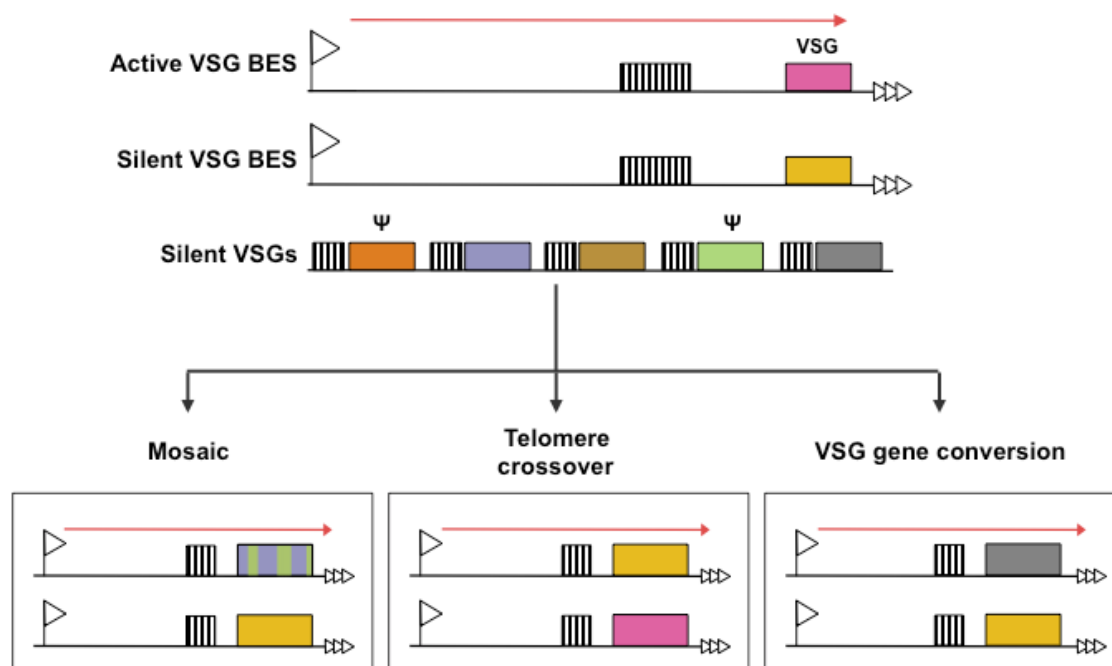
intermediate chromosomes can also provide new VSGs by recombination and appear to be preferentially used early in the course of an infection (Hall *et al.*, 2013; Morrison *et al.*, 2005).

A number of distinct reactions have been described in VSG switching by recombination: gene conversion (VSG GC), reciprocal telomere exchange (telomere crossover, telomere XO) and mosaic VSG formation (Fig. 1-5). Gene conversion (see Section 1.3.2.1 for discussion) involves the duplicative transposition of a silent, intact VSG and surrounding sequence into the active VSG BES. The boundaries of the duplicated region frequently extend beyond the VSG. Upstream, the boundary of gene conversion can be the 70 bp repeats but can also extend beyond this when the donor VSG is in a BES (Liu *et al.*, 1983; Timmers *et al.*, 1987). Downstream sequence similarity can involve conserved 3' VSG open reading frame (ORF) sequences, but often uses common sequences in the untranslated regions (UTRs) of the VSG (Bernards *et al.*, 1981) or can extend beyond to the end of the chromosome if the silent VSG is telomeric (Delange *et al.*, 1983). The latter is known as telomere conversion. Although 70 bp repeats are used for homology, they are not indispensable for gene conversion, since VSG switching events by gene conversion are still observed when these repeats are deleted from the VSG BES or are inverted (McCulloch *et al.*, 1997). This is because, at least in monomorphic lines, such gene conversion reactions often utilise sequence similarity further upstream between VSG BES (Lee & Van der Ploeg, 1987), sometimes encompassing much of the *ESAGs*, even as far as the VSG promoter ~50 kb upstream (Hertz-Fowler *et al.*, 2008). Synthesis-dependent strand annealing (SDSA) (Borst *et al.*, 1996) and break-induced replication (BIR) (Barry & McCulloch, 2001) (discussed in Sections 1.3.2.2 & 1.3.2.3 respectively) have been proposed as specific recombination reactions that could account for such gene conversion events. The involvement of BIR in *T. brucei* VSG switching was supported in a study Boothroyd *et al.* (2009) analysing switched clones (discussed in Section 1.2.4).

Reciprocal telomere exchange (also called telomere crossover, telomere XO) involves the exchange of telomeres between chromosomes by crossing over and transfers a silent telomeric VSG into the active VSG BES (Pays *et al.*, 1985), though they occur less frequently than other forms of VSG recombination and may even be suppressed (Kim & Cross, 2010; Kim & Cross, 2011). The previously

active *VSG* is concurrently moved to the other chromosome telomere and is therefore retained intact on the other chromosome. It is likely that this process is achieved by DNA double strand break repair (DSBR) (Morrison *et al.*, 2009) - see Section 1.3 for discussion of DSBR.

Mosaic *VSG* formation involves the assembly of mosaic *VSG* genes through a number of segmental gene conversion recombination reactions. Portions of two or more donor *VSG* genes or pseudogenes are combined to form a novel, functional *VSG* gene by segmental *VSG* gene conversion (Barbet *et al.*, 1982; Barbet *et al.*, 1989; Kamper & Barbet, 1992; Thon *et al.*, 1989; Thon *et al.*, 1990). Mosaic gene formation allows an expansion of the number of possible *VSG* genes beyond those already encoded intact in the genome. As with other instances of homologous recombination in *T. brucei* (Barnes & McCulloch, 2007), mosaic *VSG* formation relies on sequence similarity. However, in this case the reaction must detect some sequence similarity within the *VSG* ORFs (Marcello & Barry, 2007), perhaps reflecting the preference for mosaic *VSG* donors sharing a high degree of sequence similarity (>73%) (Hall *et al.*, 2013). In contrast to *VSG* gene conversion, the formation of mosaic *VSGs* is inefficient (Marcello & Barry, 2007), with mosaics expressed later in infections (Hall *et al.*, 2013). It is assumed that mosaic gene formation involves assembly intermediates, which has been suggested to occur in a silent *VSG* BES (Barry & McCulloch, 2001; Marcello & Barry, 2007), though assembly intermediates have not been described in these loci or elsewhere.



**Figure 1-5 Pathways of VSG switching by recombination**

As described in the text, VSG switching by recombination can occur by mosaic VSG formation, telomere crossover and VSG gene conversion. A VSG (pink box) in the active VSG BES is shown, a VSG (orange box) in a silent VSG BES and multiple silent VSGs and VSG pseudogenes ( $\Psi$ ) elsewhere in the genome. Hatched boxes depict 70 bp repeats and red arrows indicate transcription of the VSG BES. A mosaic VSG formed from multiple silent VSG and VSG pseudogene donors is depicted by the green and purple box

A number of proteins have now been identified to be involved in VSG switching by recombination. These include RAD51 (McCulloch & Barry, 1999), the RAD51-3 paralogue (Dobson *et al.*, 2011; Proudfoot & McCulloch, 2005) and BRCA2 (Hartley & McCulloch, 2008). RAD51, RAD51-3 and BRCA2 are homologous recombination (HR) components and *rad51*<sup>-/-</sup>, *rad51-3*<sup>-/-</sup> and *brca2*<sup>-/-</sup> mutants have a reduced but not abolished rate of VSG switching (Dobson *et al.*, 2011; McCulloch & Barry, 1999; Proudfoot & McCulloch, 2005). In contrast, MRE11 and the RAD51 paralogues RAD51-4, -5 and -6, which are all involved in HR, are not involved in VSG switching (Bressan *et al.*, 1999; Dobson *et al.*, 2011; Proudfoot & McCulloch, 2005; Robinson *et al.*, 1999). This shows that not all HR factors have a role in *T. brucei* VSG switching. The topoisomerase TOP3 $\alpha$  (Kim & Cross, 2010) and RMI1 (Kim & Cross, 2011), components of a complex, which in yeast acts as a regulator of crossover (Bussen *et al.*, 2007; Chelysheva *et al.*, 2008; Raynard *et al.*, 2008), are also involved in *T. brucei* VSG switching by recombination. Unlike *brca2*, *rad51* and *rad51-3* mutants described above, *topo3a*<sup>-/-</sup> and *rmi1*<sup>-/-</sup> mutants display an elevated VSG switching rate, as well as an increase in VSG switching events involving recombination at the 70 bp repeats (Kim & Cross, 2010; Kim & Cross, 2011). These studies demonstrate that factors

that act at multiple stages of HR can influence not only the rate of VSG switching, but the switching profile.

### **1.2.4 *T. brucei* 70 bp repeats**

The majority (>90%) of VSG genes in *T. brucei* have one or more 70-bp repeats upstream of them. These apparently VSG-specific repeats were first described by Liu *et al.* (1983), and sequence comparisons have shown that the repeat elements are quite variable in sequence and size between loci, though they are composed of three sections: a 5' section, containing 5-116 TAA or TAA-like tandem repeats, a moderately conserved sequence with an alternating GT characteristic, and a highly conserved 3' sequence that is mainly alternating TA repeats (Aline *et al.*, 1985). Much of the size variation is due to changes in numbers of (TAA)<sub>n</sub> repeat elements within each repeat (Aline *et al.*, 1985), though sequence composition and length changes are found in other conserved sub-elements of these complex repeats (Shah *et al.*, 1987).

In their analysis of the VSG gene silent archive, Marcello & Barry (2007) found that 92% of full-length array VSGs have at least one 70 bp repeat 1-2 kbp upstream with 74% of these possessing only one, 14% possessing two and 5% three to fifteen copies. However, the number of 70 bp repeats adjacent to VSG BES located VSGs is generally greater, comprising tens to over 100 repeats (Alsford *et al.*, 2009). MES VSGs are also flanked by 70 bp repeats, but the array lengths appear more comparable with subtelomeric VSGs (Ginger *et al.*, 2002). It has long been thought that the 70 bp repeats play an important role in VSG switching, though whether they provide passive sequence similarity or are the focus for more active promotion of recombination is not clear.

#### **1.2.4.1 DNA double strand breaks at 70 bp repeats as initiators of VSG switching**

The 70 bp repeats are a 'hot spot' for recombination events, marking the 5' border of many VSG gene conversion events (Campbell *et al.*, 1984; Liu *et al.*, 1983). This contrasts with the 3' border, which is different depending on the type of conversion event. In gene conversion of array VSGs the 3' border is frequently within the 3' UTR of the VSG (Bernards *et al.*, 1981), whereas it is the chromosome end in the case of telomere conversion (Delange *et al.*, 1983).

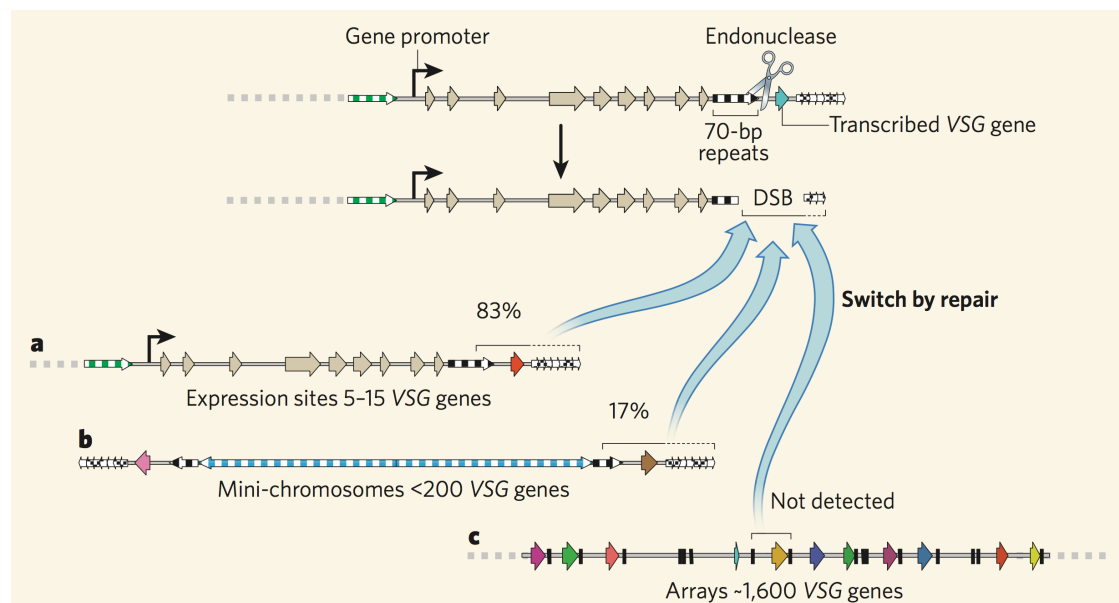
These conclusions were drawn from analysis and comparison of DNA sequencing data of the VSG and VSG BESs before and after VSG switching events. Taken together, these data may suggest that the 70 bp repeats provide 5' sequence specificity in VSG switching that is lacking at the 3' VSG end. Along with the high proportion of VSGs with 70 bp repeats upstream, in particular at VSG BESs, and the absence of significantly detectable numbers of 70 bp repeats elsewhere in the *T. brucei* genome, this has led to the theory that they may be involved in VSG switching.

Given the very limited sequence similarity between VSGs, one hypothesis is that the 70 bp repeats simply serve as stretches of sequence similarity for recombination (Delange *et al.*, 1985). Other authors have hypothesised that they serve as the initiation point for gene conversion reactions and are the location of the DNA double strand breaks (DSBs) or other lesions required for the reaction (Matthews *et al.*, 1990; Pays *et al.*, 1983). It was suggested that such initiation could depend on unusual structures formed by the 70 bp repeats that lead to cleavage within the repeat tract (Campbell *et al.*, 1984; Liu *et al.*, 1983). Alternatively, they could be recognised by a structure- or sequence-specific endonuclease (Matthews *et al.*, 1990), both of which would lead to the DSB thought necessary to initiate gene conversion. However, in experiments where the 70 bp repeats were removed from the active VSG BES, the overall VSG switching rate was not decreased and VSG gene conversion was still observed, nor was an effect observed when the repeats were mutated by inversion (McCulloch *et al.*, 1997). This suggests that the 70 bp repeats are not indispensable for VSG gene conversion reactions and it was hypothesized that mutation of the repeats made the involvement of a structure- or sequence-specific nuclease or formation of unusual structures unlikely. However, sequencing analysis of switched variants revealed that in cells lacking 70 bp repeats 22% had switched the VSG in a gene conversion reaction that utilised sequence similarity from an upstream ESAG, a pattern also observed in cells with inverted 70 bp repeats but not in wild type trypanosomes. These findings indicate that although the 70 bp repeats are not absolutely required for VSG switching, they are a crucial determinant for efficient VSG switching by gene conversion (McCulloch *et al.*, 1997).

Although DSBs were suggested to occur at 70 bp repeats, initiating gene conversion reactions and a homology search, physical evidence for this was lacking for a number of years. A study by Boothroyd *et al.* (2009) experimentally investigated this hypothesis using an inducible yeast meganuclease, I-SceI. The I-SceI meganuclease, originally isolated from *Saccharomyces. cerevisiae*, is a site-specific endonuclease that recognises and cleaves a specific 18 bp sequence. A recognition site for the yeast meganuclease was placed adjacent to the 70 bp repeats of the active VSG BES and an inducible copy of the I-SceI gene also introduced. The results of this study are summarised in Figure 1-6. The researchers reported a 250-fold increase in VSG switching frequency upon induction of a DSB, a much higher switching rate than normally found in lab-adapted *T. brucei* cell lines, and more closely resembling the switching rate of non-lab adapted trypanosomes. This increase in switching was not recapitulated when the DSB was induced within an upstream VSG pseudogene, nor after DSB induction when the 70 bp repeats were replaced with an I-SceI site. This indicates that the location of the DSB (adjacent to the 70 bp repeats) and the repeats themselves are important for the increase in VSG switching. PCR amplification and sequence analysis of 18 of the switched variants revealed that all had lost the VSG at the active VSG BES and replaced it with a donor VSG from either a silent VSG BES or a mini-chromosome. Furthermore, in all of the switched clones, the sub-telomeric region downstream of the active VSG was lost and likely replaced by the sub-telomeric region of the donor VSG. The authors conclude that such a recombination reaction is likely to be a result of BIR (discussed in Section 1.3.2.2).

Boothroyd *et al.* (2009) then investigated whether DSBs within the 70 bp repeats occur naturally *in vivo*. Using ligation-mediated PCR on unmodified, wild type *T. brucei*, they detected putative DSBs distributed over the 70 bp repeats of the active VSG BES site as well as, though less frequently, upstream of the VSG pseudogene where another shorter set of 70 bp repeats is located. Ligation-mediated PCR did not detect breaks within an internal chromosome locus indicating that DSBs occur naturally and specifically within the 70 bp repeats of wild type trypanosomes. Additionally, DSBs were not detected in a silent VSG BES, suggesting that transcription may play a role in DSB formation. The authors suggest that presence of DSBs upstream of the VSG pseudogene could be due to

DSBs occurring throughout the VSG BES but that only those in the 70 bp repeats last long enough to be detected.



**Figure 1-6 Effect of DSB induction at the 70 bp repeats on VSG switching**  
Boothroyd *et al.* (2009) used an I-SceI recognition site and enzyme to induce a DSB adjacent to the active VSG BES 70 bp repeats. Cells consequently switched VSG by recombining a silent VSG from elsewhere in the genome (other ESs, mini-chromosomes and silent VSG arrays) into the region of the DSB, replacing the previously expressed VSG. Coloured arrows indicate different VSGs, black and white boxes upstream are regions of 70 bp repeats, grey arrows are non-VSG genes. Percentages indicate the frequency of VSG switching events detected that utilised VSGs from each region. Figure from Barry & McCulloch (2009).

Further work by Glover *et al.* (2013a) supported the hypothesis that DSBs at the 70 bp repeats initiate VSG switching. Similarly to Boothroyd *et al.* (2009), Glover *et al.* (2013a) utilised the I-SceI meganuclease system to induce a DSB at one of three locations in the active BES: just downstream of the 70 bp repeats of the active VSG BES, adjacent to the BES promoter, or downstream of the VSG (proximal to the telomere). Clonal cell survival following induction of a DSB adjacent to the promoter was <20% but was strikingly lower (~5%) when a DSB was induced downstream of the 70 bp repeats or proximal to the telomere. This indicates that DSBs in the active VSG BES are typically lethal. Glover *et al.* reported that DSBs induced near the active VSG BES promoter rarely triggered VSG switching and telomere proximal DSBs induced VSG switching in 28% of surviving clones. However, almost all surviving clones switched VSG when a DSB was induced adjacent to the 70 bp repeats. Furthermore, DSBs induced adjacent to the 70 bp repeats triggered VSG switching mediated by recombination using the 70 bp repeats, indicated by retention of *ESAG1*, an



ESAG located upstream of the 70 bp repeats. In contrast, telomere proximal DSBs triggered VSG switching through loss of the VSG BES. Additionally, naturally occurring DSBs were reported in the VSG BES, but in contrast to the findings of Boothroyd *et al.* (2009), they were present at all three VSG BES regions assayed, irrespective of whether the site was transcriptionally active or inactive. Similarly, Jehi *et al.* (2014b) also detected DSBs in inactive VSG BESs, suggesting the hypothesis that the breaks exclusively localise in the inactive VSG BES and in the 70 bp repeats may be inaccurate.

The above studies demonstrate that an artificially induced DSB at the 70 bp repeats of the active VSG BES can induce VSG switching and that DSBs, or at least DNA breaks, naturally occur at the 70 bp repeats of the active VSG BES. Whether DSBs also occur within inactive VSG BESs is however, debated. Nonetheless, the current prevailing hypothesis for the initiation of VSG switching by recombination in *T. brucei* is that DSBs adjacent to the 70 bp repeats initiate VSG switching through a homology search (Boothroyd *et al.*, 2009; Glover *et al.*, 2013a).

#### **1.2.4.2 70 bp instability**

A number of events can result in DSBs and the initiation of HR, including endonuclease cleavage at specific target sites (Arazoe *et al.*, 2014; Malkova *et al.*, 2000), replication fork instability or collapse (Labib & Hodgson, 2007), unusual secondary DNA structure (Bochman *et al.*, 2012; Burrow *et al.*, 2010) and transcription (Merrih *et al.*, 2011; Takeuchi *et al.*, 2003). As such, a number of hypotheses can be proposed to explain why the 70 bp repeats may be more prone to DSBs than other sites in the *T. brucei* genome. It was first suggested a number of years ago (Matthews *et al.*, 1990) that a sequence or structure -specific endonuclease may be responsible for generating DSBs specifically within the 70 bp repeats, similar to yeast mating type switching. The *S. cerevisiae* HO endonuclease makes a site-specific DSB to initiate mating type switching that occurs by HR, reviewed by Haber (1998). However, such an endonuclease has yet to be identified.

A number of alternative explanations exist for the occurrence of DSBs at 70 bp repeats based on the repetitive nature of the repeats. The 70 bp repeats

contain a TAA·TTA motif, known to form a non H-bonded structure, which under certain conditions can destabilise the DNA helix - a property not shared by other triplet repeats (Ohshima *et al.*, 1996b). Several different triplet repeats are well established to cause repeat expansion and, when within or proximal to disease genes, genetic mutation resulting in disease in humans (Ohshima *et al.*, 1996a). Such triplicate repeats are also associated with DNA polymerase pausing (Kang *et al.*, 1995). For example, DNA polymerase pausing occurs at the CTG·CAG triplet repeat whose expansion is associated with myotonic dystrophy (Kang *et al.*, 1995), and is a hotspot for recombination (Jakupciak & Wells, 1999; Pluciennik *et al.*, 2002). Pausing of DNA polymerase on triplet repeats, including the CTG·CAG repeat, leads to replication slippage and rearrangement (Viguera *et al.*, 2001). However, the TAA·TTA repeat does not share similar properties to these repeats; it does not exhibit DNA polymerase pausing and overall has very different genetic and biochemical properties (Ohshima *et al.*, 1996b), suggesting that its role in gene regulation may differ from other triplet repeats. Transcription of the highly repetitive 70 bp repeats in the active VSG BES, which have a propensity to form this unstable structure, may destabilise the active VSG BES leading to a DSB (Boothroyd *et al.*, 2009).

A further hypothesis proposes that the 70 bp repeat region is a 'fragile site' due to the repetitive sequences within the region and thus pre-disposed to replication stalling. Replication fork stalling is associated with a number of elongation-blocking lesions including DNA damage, protein, transcription, DNA secondary structure or DNA repeats (Labib & Hodgson, 2007; Merrikh *et al.*, 2011). Replication fork collapse results in single stranded DNA (ssDNA) gaps, which after processing could become structures able to undergo HR and potentially initiate VSG switching. There are multiple ways a replication fork could stall in a VSG BES. As already mentioned, TAA·TTA repeats present in the 70 bp repeats predispose the region to forming unusual DNA structures. Such structures, or indeed the repeats themselves, could lead to replication fork stalling, as has been observed at some sites (e.g. long genes) in other eukaryotes (Brewer, 1988; Helmrich *et al.*, 2011; Tuduri *et al.*, 2009). This suggestion fits well with the findings by Boothroyd *et al.* (2009) who only detected DSBs in an actively transcribed VSG BES, though does not explain the presence of DSBs in inactive VSG BESs reported by others (Glover *et al.*, 2013a; Jehi *et al.*, 2014b).

Furthermore, replication stalling has been documented to initiate gene rearrangement. For example, the induction of a DSB in the *S. cerevisiae mat1* locus is achieved through replication stalling, which leads to recombination and gene transposition, resulting in mating type switching (Klar, 2007). Another interesting mechanism by which replication stalling may occur is demonstrated in the antigenic variation of the *Neisseria PilE* locus. The formation of a DNA:RNA hybrid, brought about by expression of a small RNA (sRNA), is necessary for the formation of a guanine quadruplex (G4) secondary DNA structure required for switching of the downstream *PilE* locus (Cahoon & Seifert, 2009; Cahoon & Seifert, 2013). G4 structures are guanine-rich regions of DNA capable of forming a four-stranded DNA structure (Huppert, 2008). It is thought that the RNA:DNA hybrid forms on the C-rich strand, allowing the formation of the G4 structure on the G-rich strand, possibly mediated by binding proteins. G4 structures can impair replication, leading to replication instability (Lopes *et al.*, 2011; Rizzo *et al.*, 2009; Weitzmann *et al.*, 1996).

### 1.3 Double Strand Break Repair (DSBR)

The genome is subjected to various types of DNA damage including DSBs, which are frequently described (Jackson, 2002; Lees-Miller & Meek, 2003b; Polo & Jackson, 2011; Zhang *et al.*, 2009) as the most serious form of DNA damage due to loss of integrity of both strands. DSBs can be caused by a number of events, including UV (Greinert *et al.*, 2012) and ionising radiation (Costes *et al.*, 2010), replication stalling (Unno *et al.*, 2013), mutagenic chemicals (Driessens *et al.*, 2009; Ma *et al.*, 2011; Morel *et al.*, 2008) and endogenously generated free radicals (Albino *et al.*, 2006). DSBs that arise must be repaired to prevent serious mutations such as gross chromosomal rearrangements, which are detrimental and possibly lethal to single-celled organisms. *T. brucei*, like all eukaryotes, possess multiple mechanisms to repair DSBs. Non-homologous end joining (NHEJ) and HR are the two principle pathways used by most eukaryotes to repair DSBs. These two pathways differ in their requirement for a homologous template to repair DSBs and in the fidelity of the repair they catalyse. Both pathways are found in almost all eukaryotes, but whether they both act in *T. brucei* is unclear.

### 1.3.1 Non-homologous end-joining (NHEJ)

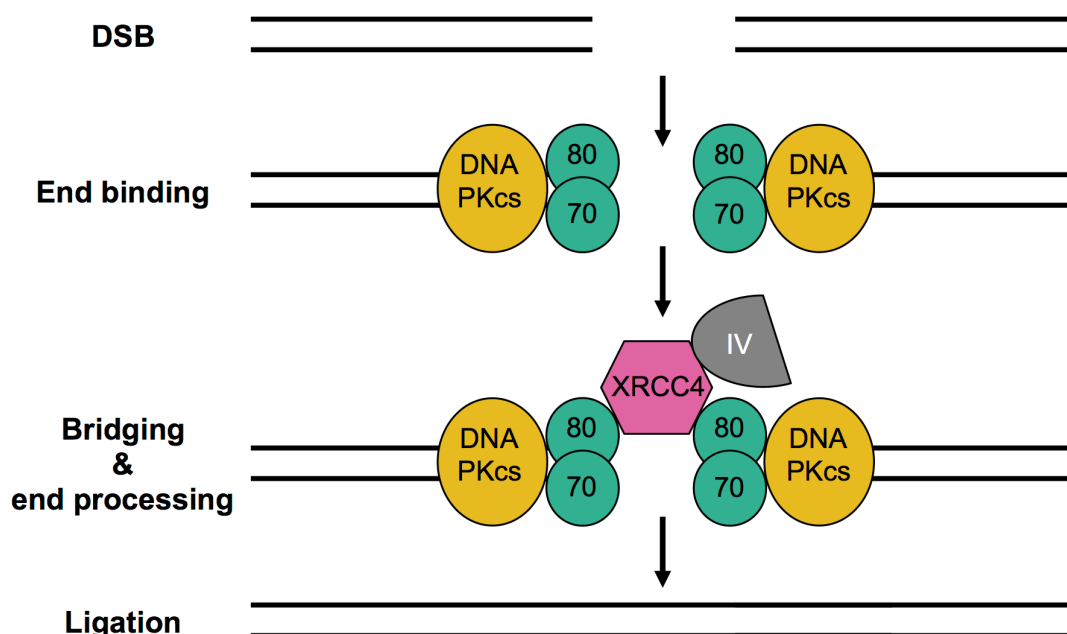
Non-homologous end joining (NHEJ) is a DSB repair pathway that is conserved in virtually all eukaryotes and is also found in some bacteria (Weller *et al.*, 2002). The pathway requires no template and little (2 or 3 bp) or no homology for repair, and thus can lead to mutations through the re-joining of resected DNA at DSB ends or the re-joining of incorrect DSB ends. Despite this, evidence indicates that most mammalian cells utilise NHEJ more than HR, as NHEJ is favoured in G0 and G1 cell cycle stages (Lees-Miller & Meek, 2003b). NHEJ is also favoured when DSBs are located proximal to the telomere (Ricchetti *et al.*, 2003). The regulation of pathway choice (HR vs NHEJ) is reviewed in Symington & Gautier (2011) and Chapman *et al.* (2012).

The primary proteins responsible for NHEJ (Fig. 1-7) are the Ku70/80 heterodimer and DNA ligase IV, which acts in association with the protein XRCC4 (Grawunder *et al.*, 1997; McElhinny *et al.*, 2000). Beyond these core proteins, NHEJ in mammals also requires the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) (Smith & Jackson, 1999), while at least in yeast, the Rad50-Mre11-Xrs2 (MRX) complex, an endonuclease, is necessary (Chen *et al.*, 2001a).

The first step of NHEJ is DSB end binding. The Ku70/80 heterodimer binds the DSB ends and can be translocated from the DNA ends in an ATP dependent manner (Dylan & Yoo, 1998). In mammals, this is thought to be important for stabilising and recruiting the DNA-PKcs, which also binds to each side of the break; DNA-PKcs is recruited by Ku70/80 and forms the DNA-PK complex (Dylan & Yoo, 1998). In yeast, MRX (MRN, Mre11-Rad50-Nbs1 in humans) processes the DSB ends (Khanna & Jackson, 2001; Paull & Gellert, 1998). This role appears to have been supplanted by DNA-PKcs in mammals. Some bacteria, such as *Bacillus subtilis*, possess a NHEJ pathway and Ku homologues have been found, suggesting that the pathway arose before the divergence of the prokaryotic and eukaryotic lineages (Weller *et al.*, 2002). The molecules of DNA-PKcs either side of the break bridge the ends together prior to re-ligation (DeFazio *et al.*, 2002).

The final stage of NHEJ is ligation. In eukaryotes, the NHEJ-specific DNA ligase IV/XRCC4 complex is recruited to the DNA ends and catalyses the ligation of the

two ends (Chen *et al.*, 2000). This process often results in loss of nucleotides at the two ends of the break, resulting in inaccurate repair.



**Figure 1-7 Schematic of the steps of non-homologous end joining**  
The steps of NHEJ are shown, as described in the text. Black lines represent duplex DNA. Ku70/80 heterodimers bind the DNA ends (turquoise circles), recruiting DNA PKcs (yellow oval) to the DSB ends. The DNA ligase IV-XRCC4 complex (grey semi-circle and pink hexagon respectively) completes the ligation of the DSB ends. Figure adapted from Claire Hartley, PhD Thesis (2008).

No evidence has so far been found that NHEJ takes place in *T. brucei*; the genome does not encode a detectable DNA ligase IV-XRCC4 complex, though Ku70/80 is conserved (Burton *et al.*, 2007; Conway *et al.*, 2002a). In addition, transformation experiments (Conway *et al.*, 2002b), induction of a DSB (Glover *et al.*, 2011; Glover *et al.*, 2008) and analysis of end joining in cell extracts (Burton *et al.*, 2007) has each failed to provide evidence for NHEJ. *T. brucei* does however possess an ‘MMEJ’ repair system (microhomology-mediated end joining, see below) capable of joining the free ends that is readily seen in each of the above experimental settings, perhaps suggesting that it has evolved to supplant NHEJ.

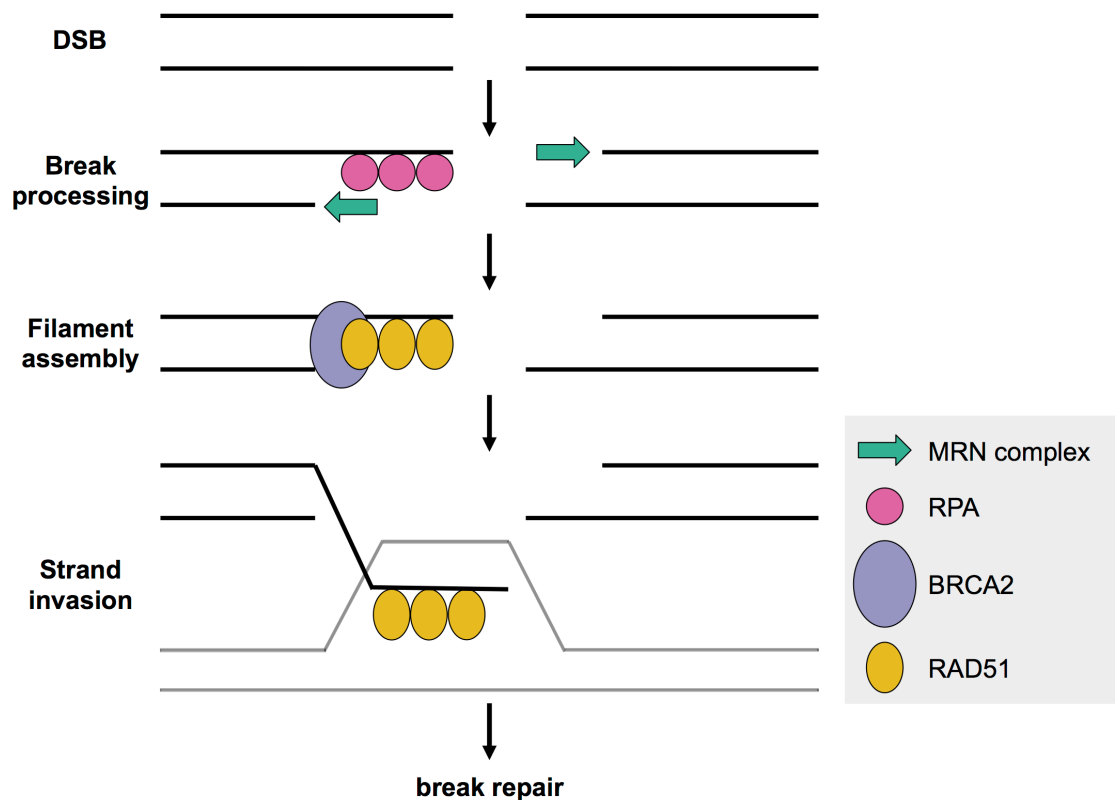
### 1.3.2 Homologous recombination (HR)

DSB repair by homologous recombination (HR) involves the utilisation of an intact, sequence-related DNA molecule located elsewhere in the genome as a template for repair. For this reason, HR normally results in fewer mutations

compared to NHEJ. In contrast to NHEJ, which predominantly acts in G0/G1 (Lees-Miller & Meek, 2003a), HR acts primarily in late S-phase/G2 in vertebrate cells (Takata *et al.*, 1998), reviewed in Branzei & Folani (2008). HR can be considered to occur in stages, as shown in Figure 1-8.

The first stage of HR, termed synapsis, involves the invasion of one strand of the broken DNA duplex into a homologous sequence, normally an intact duplex, elsewhere in the genome. This requires a 3' ssDNA overhang coated with a filament of the protein RAD51, a eukaryotic homologue of the bacterial RecA protein and a key protein in HR, permitting strand exchange to take place. The 5'-3' exonuclease Exo1, MRX (Mre11-Rad50-Xrs2) complex and Sae2 resect the 5' DNA ends to produce 3' ssDNA overhangs necessary for binding of Rad51 (Cannavo *et al.*, 2013; Fiorentini *et al.*, 1997; Nicolette *et al.*, 2010). The ssDNA in eukaryotes is coated with replication protein A (RPA), which protects the ssDNA from nucleases and removes secondary structure but is also inhibitory to nucleoprotein filament formation (Sugiyama *et al.*, 1997). RPA removal is therefore required for RAD51 filament formation. The RAD52 recombinase, which is not detectably present in *T. brucei*, facilitates this in *S. cerevisiae* (Sugiyama & Kowalczykowski, 2002; Sung, 1997) and BRCA2 has been demonstrated to bind RAD51 and promote RAD51 filament formation (Jensen *et al.*, 2010; Liu *et al.*, 2010b; Thorslund *et al.*, 2010).

Following assembly of the RAD51-nucleoprotein filament, strand exchange occurs. RAD51 catalyses the 3' ssDNA end invasion of an undamaged, homologous DNA molecule forming a three-way junction (D-loop). Though RAD51 is the catalytic enzyme in strand invasion, it is supported by other proteins including RAD52, RAD54, so-called 'RAD51 paralogues' (such as RAD55 and RAD57 in yeast) and BRCA2 (Llorente *et al.*, 2008). *T. brucei* encodes four Rad51 paralogues that act in HR (Dobson *et al.*, 2011; Proudfoot & McCulloch, 2005). DNA synthesis then occurs using the D-loop 3' DSB end as the primer and the homologous DNA strand as the template (Paques & Haber, 1999). After priming DNA synthesis on the 3' DSB end, repair of the DSB can proceed in one of several pathways: gene conversion, synthesis dependent strand annealing (SDSA) or BIR.



**Figure 1-8 Schematic of homologous recombination in eukaryotes**  
 The stages of HR, as described in the text. Black lines represent duplex DNA and grey lines represent duplex DNA used as the template for repair. The 3' ends of the DSB are resected by the MRN complex (turquoise arrow) and RPA (pink circles) binds the ssDNA. RAD51 (orange ovals) is loaded onto the 3' ssDNA by BRCA2 (purple oval) and the RAD51 filament invades a homologous DNA template (grey line).

### 1.3.2.1 Gene conversion

Gene conversion is the main route of DSB repair by HR (Chen *et al.*, 2007) and allows the unidirectional transfer of genetic information from homologous sequences to the site of the DSB. Gene conversion can occur via the formation of a Holliday junction (HJ) (Fig. 1-9), or alternatively by synthesis-dependent strand annealing (SDSA) or BIR (not depicted, described below). In both pathways, the 3' DSB end invades the homologous sequence forming a D-loop. Gene conversion by HJ resolution then proceeds by the other 3' ssDNA tail pairing with the D-loop, resulting in the formation of two HJs (Holliday, 1964). This intermediate is resolved by cleavage of the Holliday junctions to yield either a product with or without crossover of flanking regions, depending on the axis along which the HJ is cleaved. Gene conversion by SDSA produces non-crossover products (Sung & Klein, 2006).

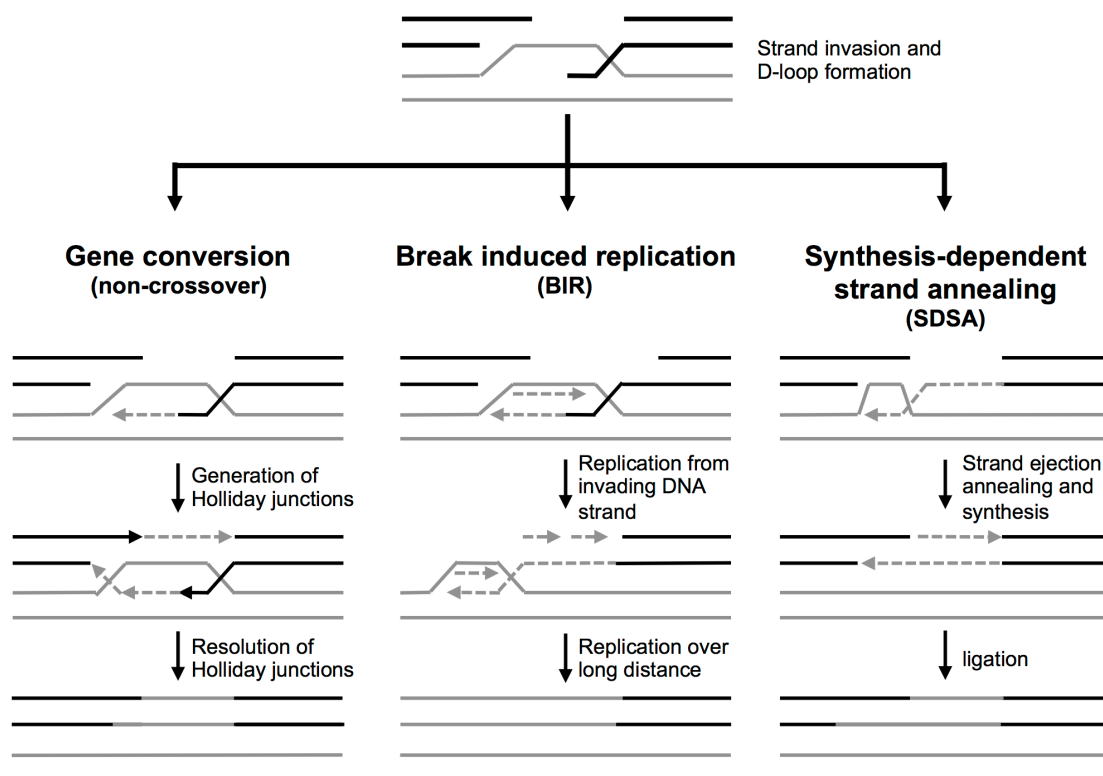
### 1.3.2.2 Break-induced replication (BIR)

Break-induced replication (BIR) (Fig. 1-9) occurs when only one of the 3' ssDNA strands invades a homologous DNA duplex, forming a D-loop. BIR can be used to repair DSBs when only one end is available for strand invasion and D-loop formation, such as when homology only exists at one end of the DSB (Llorente *et al.*, 2008). The D-loop assembles into a replication fork and copies the downstream sequence, leading to long gene conversion tracts that can be hundreds of kilobases in length or even to the chromosome end (Esposito, 1978; Malkova *et al.*, 1996; Paques & Haber, 1999). The frequent and long gene conversions cause BIR reactions to often go undetected in cells (Paques & Haber, 1999). Such gene conversion events differ from SDSA (see Section 1.2.2.3) and DSB repair because resolution is by establishment of a bubble-like replication fork that moves along the chromosome (Saini *et al.*, 2013) rather than through a process driven by recombination. BIR can occur through Rad51-dependent and Rad51-independent mechanisms (Malkova *et al.*, 1996; McEachern & Haber, 2006); Rad51-independent BIR requires shorter homologous sequences (30 bp) than Rad51-dependent BIR (Ira & Haber, 2002).

### 1.3.2.3 Synthesis-dependent strand annealing (SDSA)

In contrast to BIR and gene conversion, in synthesis-dependent strand annealing (SDSA) the D-loop migrates and the invading DNA strand (forming the D-loop) is ejected from the D-loop after priming DNA synthesis. The ejected strand anneals with the complementary strand associated with the other end of the DSB (Nassif *et al.*, 1994) (Fig. 1-9). DNA synthesis and ligation complete the repair. As no HJ is formed, only non-crossover products result from SDSA (San Filippo *et al.*, 2008).



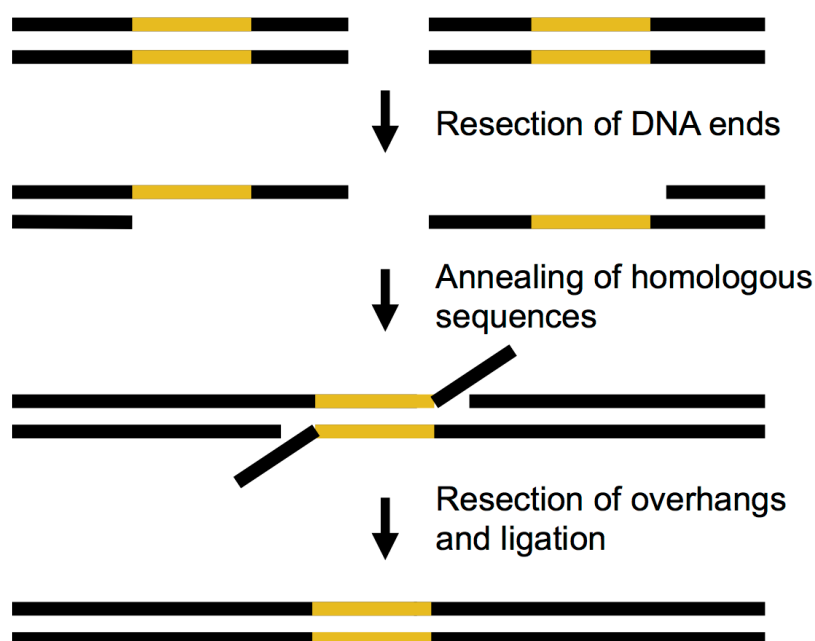


**Figure 1-9 Pathways of DSB repair**

Three mechanisms, described in the text, used to repair DSBs by homologous recombination are shown. Solid lines indicate duplex DNA, grey lines represent template DNA and dashed lines indicate newly replicated DNA. Top: a DSB is shown that has been processed and one ssDNA end has invaded a homologous template. Repair of the DSB by gene conversion, BIR and SDSA are shown.

### 1.3.3 Single strand annealing (SSA)

A fourth mechanism, single strand annealing (SSA) (Fig. 1-10), can be employed to repair DSBs. It occurs when a DSB is flanked by direct repeats and involves repair of the DNA using the repeated DNA sequences. SSA is independent of Rad51 (Ivanov *et al.*, 1996), does not involve the strand invasion and D-loop formation described in the mechanisms above and does not produce HJs. Instead, following resection of the DSB ends to produce 3' ssDNA, the 3' DSB ends anneal to one another rather than invading a homologous DNA sequence. The Rad52 recombinase is required for SSA (Ivanov *et al.*, 1996). Glover & Horn (2009) found that repair of DSBs within the *RRNA* spacer locus, which is flanked by stretches of duplicated sequence, was consistent with SSA, suggesting this process operates in *T. brucei*.

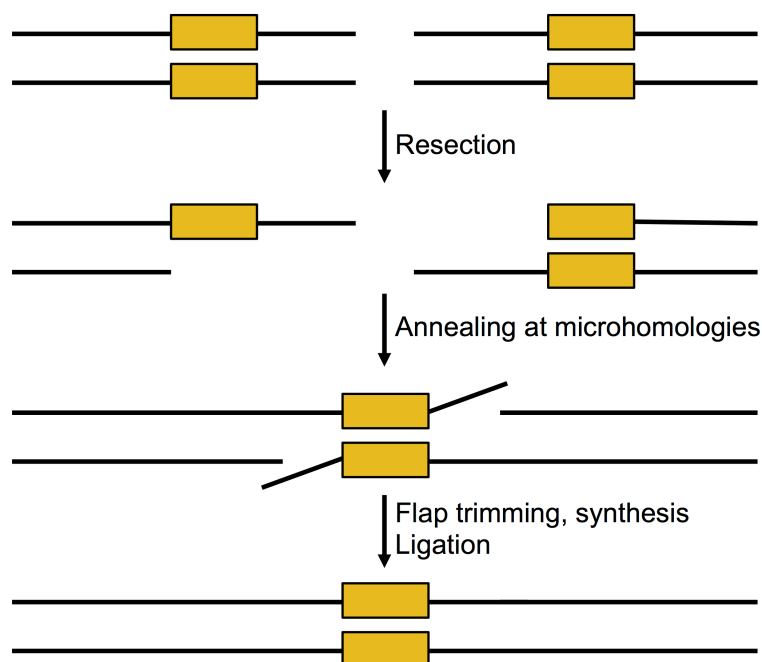


**Figure 1-10 DSB repair by single strand annealing (SSR)**  
 Duplex DNA is represented by black lines and direct repeat sequences are indicated in yellow. Resection of the non-homologous 5' ends of the DSB is followed by annealing of the direct repeats. The overhangs are resected and ligation completes repair of the DSB, resulting in loss of the sequence between the direct repeats.

### 1.3.4 Microhomology-mediated end joining (MMEJ)

Microhomology-mediated end joining (MMEJ) (Fig. 1-11) is another DSB repair pathway, potentially distinct from NHEJ and HR, though the enzymatic machinery remains unknown. Although it was initially thought that MMEJ functioned only in the absence of NHEJ machinery, it has been established that MMEJ occurs even when the NHEJ pathway is not compromised (McVey & Lee, 2008; Nussenzweig & Nussenzweig, 2007). In contrast to NHEJ, which requires no homology to repair DSBs, MMEJ requires short stretches of homology, typically ~6-20 bases, though (as with NHEJ) it is error prone and loss of sequence occurs (McVey & Lee, 2008). It has been proposed (McVey & Lee, 2008) that MMEJ repair begins with the removal of Ku70/80 and Rad51 (which inhibit MMEJ) from the DNA, followed by resection of the 5' DNA ends by the MRX complex and the nuclease Exo1 and endonuclease Sae2. This promotes resection to reveal microhomologous sequences. These micro-homologous sequences anneal to one another and repair is completed by DNA synthesis and ligation. The enzymes responsible for the short homology base-pairing are unclear and the process can be very imprecise, leading to repair products containing numerous base mismatches. MMEJ is readily detected in *T. brucei* *in vivo* (Conway *et al.*,

2002b; Glover *et al.*, 2011) and *in vitro* using nuclear cell extracts (Burton *et al.*, 2007), and may therefore be a robust process in *T. brucei*.



**Figure 1-11 Microhomology-mediated end joining (MMEJ)**  
 Duplex DNA is represented by black lines and yellow boxes indicate microhomologous sequences. After removal of Ku70/80 and RAD51, repair by MMEJ proceeds by resection of the 5' DNA ends to reveal 3' ssDNA and microhomologous sequence. These microhomologies anneal and repair is completed by trimming of 3' ssDNA flaps, gap filling DNA synthesis and ligation of ends to produce dsDNA, with loss of sequence between the microhomologies.

## 1.4 DNA Helicases in DNA repair

Helicases are molecular motors that unwind polynucleic acids using energy released from nucleoside triphosphate (NTP) hydrolysis, most commonly ATP. This process results in the translocation of the helicase along the polynucleic acid in a 5'→3' or 3'→5' direction. Helicases are divided into two major superfamilies: superfamily 1 (SF1) and superfamily 2 (SF2), along with the smaller superfamilies 3 to 6. Different types of helicase act on different nucleic acid substrates. DNA helicases unwind DNA, RNA helicases unwind RNA and some helicases unwind DNA:RNA hybrids. DNA helicases function in a wide range of nucleic acid processes. Some examples of roles helicases play in DNA replication, transcription, repair and recombination are described below.

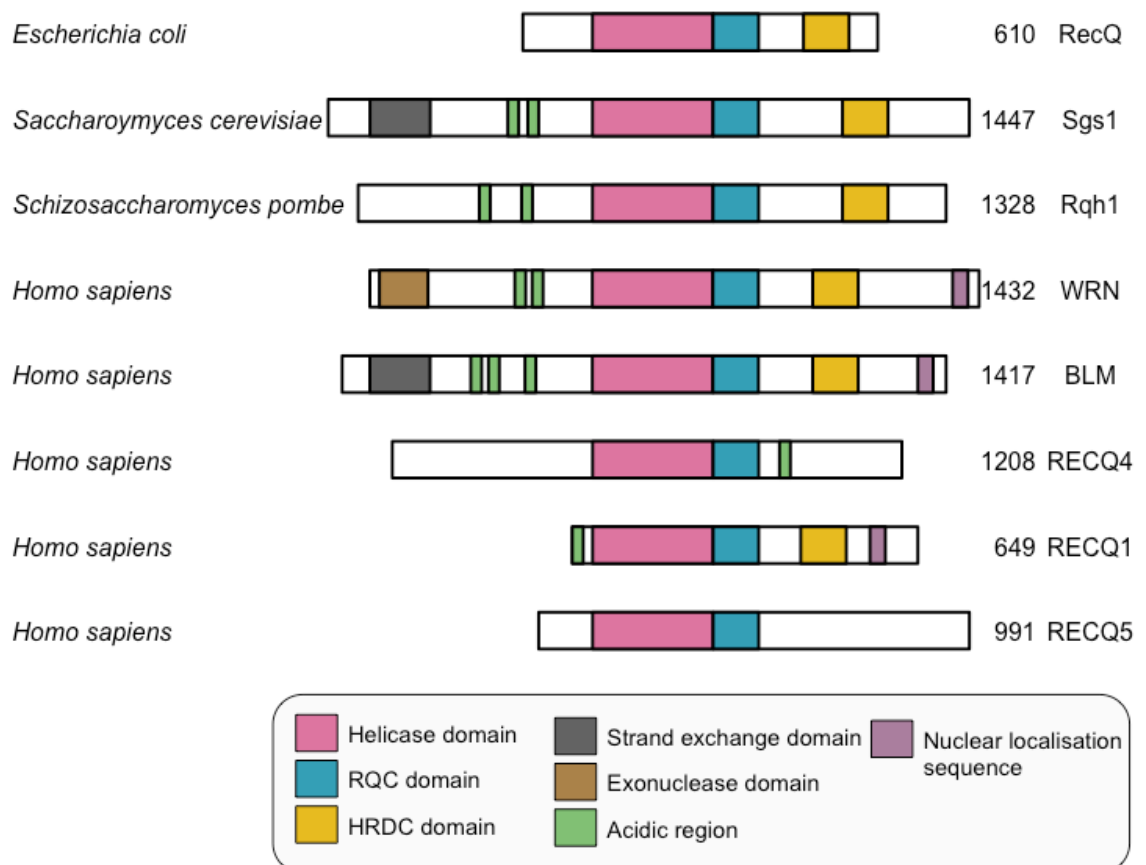
Helicases are crucial for the replication of DNA. For example, the MCM2-3 (mini chromosome maintenance complex) helicase is the eukaryotic replicative helicase (Labib *et al.*, 2000). Other helicases facilitate replication of specific regions, such as the role of the Srs2 helicase in replication of hairpins (Anand *et al.*, 2012). The recovery of replication forks that have stalled, reversed or collapsed also require helicases. For example, the Werner helicase, along with the Dna2 nuclease, functions in the recovery and restart of reversed replication forks in human cells (Thangavel *et al.*, 2015).

Transcription and transcriptional regulation also utilises helicases. The TFIIH (transcription factor II human) complex, which is part of the transcription initiation apparatus, also functions in DNA repair, and contains two DNA helicases, XPD and XPB (xeroderma pigmentosum group D and xeroderma pigmentosum type B) - reviewed in Egly & Coin (2011). Another family of helicases with important roles in transcription is the DExD/H box RNA helicases. These helicases regulate transcription in a variety of ways, reviewed in Fuller-Pace (2006), including acting as transcriptional co-activators or repressors, acting as bridging factors to recruit transcription factors and stabilising transcription initiation complexes.

XPB and XPD helicases, as part of the TFIIH complex, also function in nucleotide excision repair (NER) (Coin *et al.*, 2004; Kuper *et al.*, 2014; Timmins & McSweeney, 2006). NER is a DNA repair pathway that repairs single-stranded DNA damage, such as UV damage, which distorts the DNA helix, reviewed in Kamileri *et al.* (2012). Helicases also participate in other DNA repair pathways. Human RECQL5 (RecQ-like 5) is involved in base excision repair (BER) (Tadokoro *et al.*, 2012), a repair pathway that corrects lesions resulting processes such as alkylation, oxidation and deamination, reviewed in Robertson *et al.* (2009). In addition, a number of helicases including Sgs1/BLM and human helicase B are involved in DNA repair by homologous recombination (Liu *et al.*, 2015; Wu & Hickson, 2003). Helicases can play both stimulatory and inhibitory roles in homologous recombination. For instance, the Sgs1 helicase in yeast (Bloom helicase in humans) is part of the machinery that resects DNA ends, a necessary step for HR (Mimitou & Symington, 2008; Nimonkar *et al.*, 2011). In contrast, the Srs2 helicase has an anti-recombinogenic effect due to its disruption of RAD51-ssDNA filaments (Krejci *et al.*, 2003; Veaute *et al.*, 2003).

### 1.4.1 RecQ helicases in DNA repair and two putative *T. brucei* RecQ-like helicases

RecQ-like helicases are an SF2 family (Gorbalenya & Koonin, 1993) of DNA helicases so-called due to the homologous region they share with the *E. coli* RecQ helicase, first described by Umezu *et al.*, (1990). The RecQ family is conserved from bacteria to humans (Fig. 1-12) and unwind DNA in a 3'→5' direction (Bachrati & Hickson, 2008). *Schizosaccharomyces pombe* and *S. cerevisiae* each possess a single RecQ helicase, Rqh1 and Sgs1 respectively. In contrast, multicellular eukaryotes such as humans and *Drosophila* possess multiple RecQ helicases (Adams *et al.*, 2003; Bernstein *et al.*, 2010; Capp *et al.*, 2009; Chen *et al.*, 2010).



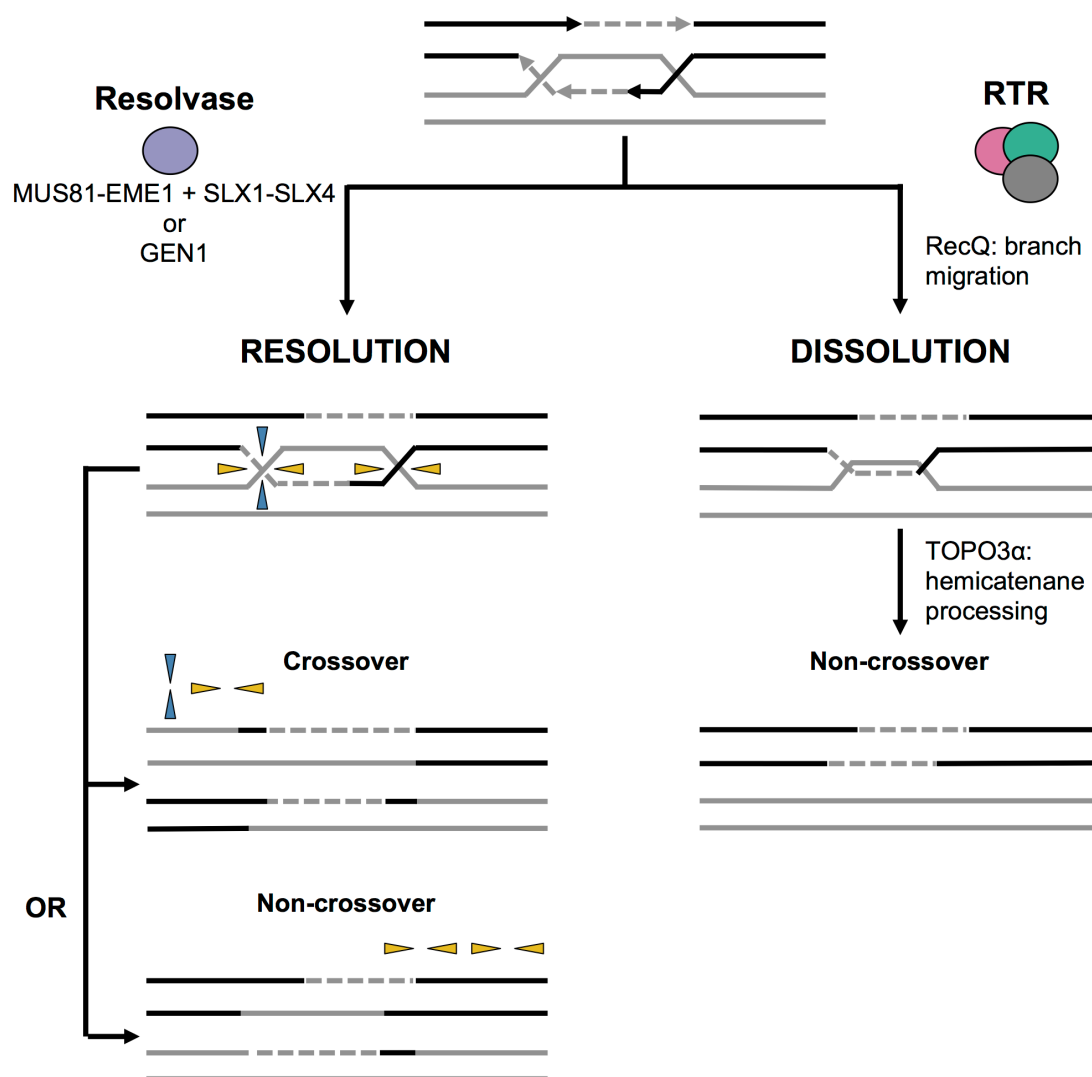
**Figure 1-12 Domains and features of bacterial and eukaryotic RecQ helicases**  
 Typical domains and features found in RecQ helicases are shown. All RecQ helicases share a conserved helicase (DEAD/DEAH box helicase) domain. RQC (RecQ C-terminal) and HRDC (helicase and RNase D C-terminal) domains are present in most RecQ family members and mediate protein and nucleic acid interactions. Some members contain a strand exchange domain, exonuclease domain or nuclear localisation signal. The number of amino acids in each protein sequence is indicated on the right. Accession numbers can be found in Appendix 7.2. Not to scale.

RecQ helicases participate in DNA replication, recombination and repair (Croteau *et al.*, 2014). RecQ proteins interact with many DNA repair proteins to play multiple roles in multiple types and stages of DNA repair. They also participate in DSB NHEJ, where the Werner syndrome RecQ protein (WRN) interacts with the XRCC4-ligase IV complex (Kusumoto *et al.*, 2008). In addition, human RECQ1 interacts with Ku70/80, with depletion of RECQ1 leading to reduced NHEJ (Parvathaneni *et al.*, 2013). HR involves RecQ proteins such as the yeast Sgs1 and human Bloom syndrome RecQ protein (BLM) helicases, which promote resection of DNA ends (Gravel *et al.*, 2008; Mimitou & Symington, 2008; Zhu *et al.*, 2008). Yeast Sgs1 also interacts with Rad51 to inhibit Rad51 filament formation (Wu *et al.*, 2001). *S. cerevisiae* Sgs1 interacts with the HR nuclease Mre11 (Chiolo *et al.*, 2005). Additionally, Sgs1/BLM interact with mismatch repair proteins Mlh1 and Mlh3 (Argueso *et al.*, 2003; Dherin *et al.*, 2009; Langland *et al.*, 2001; Wang & Kung, 2002). The function of these RecQ-mismatch repair interactions is unclear, though it has been suggested it could be to promote resolution of meiotic crossovers (Wang & Kung, 2002).

RecQ helicases also interact with many proteins involved in DNA replication. For example, human RECQ4 interacts with the replisome factors MCM10 and MCM2-7 helicase and RECQL4 interacts with CDC45 and GINS (Go, Ichi, Nii, and San; five, one, two, and three in Japanese) (Im *et al.*, 2009; Xu *et al.*, 2009b). Due to their direct role in DNA replication, absence of RecQ helicases can lead to replicative damage. For example, WRN co-purifies with the DNA replication complex in embryonic stem cells (Lebel *et al.*, 1999), and absence of BLM and WRN leads to accumulation of replication intermediates (Lonn *et al.*, 1990). RecQ helicases also act in the repair of replicative damage, as in the case of yeast Sgs1, a function that interestingly can be uncoupled from its function in DNA recombination (Bernstein *et al.*, 2009).

RecQ helicases play an important role in the processing of double Holliday junctions (dHJs) as part of the RTR (also known as BTR) complex. The complex is composed of a RecQ helicase (BLM in mammals, SGS1 in yeast), a topoisomerase III $\alpha$  and RMI1 (RecQ mediated genome instability 1; also known as BLAP75/18) (Johnson *et al.*, 2000; Wu *et al.*, 2000; Yin *et al.*, 2005). The RTR complex is an important regulator in the processing of dHJ recombination intermediates, to produce exclusively non-crossover products, known as dHJ

dissolution (Bussen *et al.*, 2007; Chelysheva *et al.*, 2008; Raynard *et al.*, 2008) (Wu & Hickson, 2003) (Fig. 1-13). This in contrast to the alternative pathway of processing dHJs by resolvases (such as Mus81), which can produce both non-crossover and crossover products, known as dHJ resolution. In dissolution, the RecQ helicase in the RTR complex branch migrates the two HJs closer to one another to form a hemicatenane intermediate, a structure in which one strand of a DNA duplex is wound around a strand of another DNA duplex. TOP3 $\alpha$  then processes the hemicatenane, making a nick in one strand and passing the other strand through to dissolve the dHJ (Yang *et al.*, 2010). RMI1 possesses no catalytic activity but its presence is essential for the stability of the complex and loss of RMI1 results in an increase in crossover events similar to that observed after BLM/SGS1 loss (Chang *et al.*, 2005; Yin *et al.*, 2005). Alternatively, resolvases can process dHJs by cleaving the dHJs and depending on the the orientation of the nucleolytic cleavage the result is either a crossover or non-crossover product. In humans this can be achieved with either the MUS81-EME1 (MUS81-MMS4 in *S. cerevisiae*) complex acting together with SLX1-SLX2, or by GEN1 (YEN1 in *S. cerevisiae*). As crossovers are potentially deleterious to the cell, the processing of dHJs by resolution is tightly regulated during the cell cycle. Dissolution via the RTR complex is favoured throughout most of the cell cycle, with resolvase activity limited to the end of the cell cycle to resolve any remaining recombination intermediates that would be a barrier to chromosome segregation (West *et al.*, 2015).



**Figure 1-13 Resolution and dissolution of double Holliday junctions**  
 A double Holliday junction (dHJ) is shown following resection of a DSB, D-loop formation and pairing of the other 3' ssDNA tail to form a dHJ. Processing by resolution involves a resolvase such as the MUS81-EME1/SLX1-SLX4 complex or GEN1 and can result in crossover or non-crossover products depending on the orientation of cleavage (coloured arrows). Alternatively, the RTR can dissolve the dHJs using the branch migration activity of the RecQ-like helicase and the decatenating activity of the topoisomerase, producing only non-crossover products.

*Trypanosoma brucei* is unusual among unicellular organisms with respect to its RecQ helicases. The *T. brucei* genome contains two genes annotated as putative RecQ-like helicases: Tb427.06.3580 (here referred to as RECQ1) and Tb427.08.6690 (here referred to as RECQ2) (TriTrypDB.org). These two *T. brucei* putative RecQ helicases are two of the factors examined in this thesis, neither having been characterised prior to this work. Most multicellular organisms possess multiple RecQ helicases (discussed below). Humans encode five: WRN (Werner syndrome protein), BLM (Bloom syndrome protein), RECQ1, RECQ4 and RECQ5 (Bernstein *et al.*, 2010). *C. elegans* encode four, named HIM-6, WRN-1,

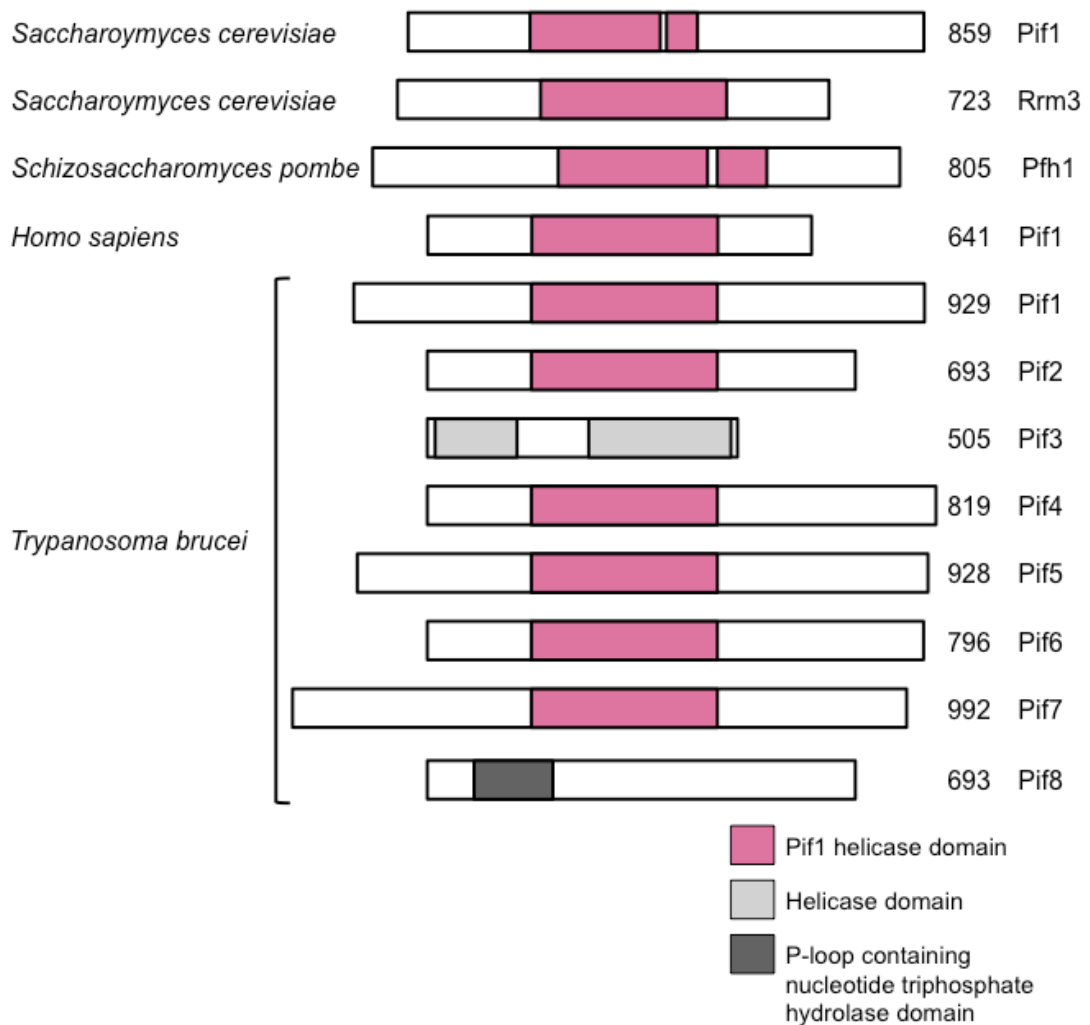


RECQ5 and RECQ1, though only the first three have been functionally characterised (Hyun *et al.*, 2012; Jeong *et al.*, 2003; Jung *et al.*, 2014). In contrast, unicellular organisms generally only contain one RecQ helicase: e.g. RecQ in *E. coli*, Rqh1 in *S. pombe* and Sgs1 in *S. cerevisiae* (Bachrati & Hickson, 2008; Myung *et al.*, 2001; Umezu *et al.*, 1990). *T. brucei* is therefore unusual among unicellular organisms in containing more than one RecQ helicase, though this would appear to be common among kinetoplastids, since searches conducted on the TriTryp database (TriTrypDB.org) show that most *Trypanosoma* and *Leishmania spp.*, as well as *Crithidia fasciculata*, contain more than one gene annotated as a RecQ-like helicase.

Kinetoplastids have evolved expanded gene families to carry out kinetoplastid DNA (kDNA) repair and replication. For example, the *Pif1* helicase gene family, restricted to one or two genes in most eukaryotes, is expanded in *T. brucei* to eight genes, six of which localise to the mitochondria and three of which have been demonstrated to function in kDNA replication or maintenance (Liu *et al.*, 2009a; Liu *et al.*, 2009b; Liu *et al.*, 2010a; Wang *et al.*, 2012). Kinetoplastids also contain an expansion in bacterial-like DNA Polymerase I-related enzymes that all act in kDNA replication or maintenance (Bruhn *et al.*, 2010; Chandler *et al.*, 2008; Concepcion-Acevedo *et al.*, 2012; Klingbeil *et al.*, 2002). Given this, it is possible that TbRECQ1 has evolved to provide a specialised kDNA function.

### 1.4.2 Pif1 family helicases in DNA repair

First described in *S. cerevisiae* (Foury & Kolodynski, 1983), and since shown to be present in almost all eukaryotes, Pif1 family helicases (Fig. 1-14) belong to the SF1B helicase family and move in a 5'→3' direction (Bochman *et al.*, 2010). *S. cerevisiae* possesses two Pif1 helicases, Pif1 and Rrm3, though most eukaryotes studied to date possess only a single Pif1 helicase (Bochman *et al.*, 2010). The founding member of the Pif1 helicase family is *S. cerevisiae* Pif1 (ScPif1), which was discovered in a genetic screen for genes affecting mitochondrial DNA (Foury & Kolodynski, 1983).



**Figure 1-14 Pif1 family helicases**

**Pif1 helicase domains in various eukaryotic Pif1 helicases and predicted protein domains in all eight *T. brucei* PIF1 helicases. The number of amino acids in each sequence is shown on the right. Accession numbers can be found in Appendix 7.2. Not to scale.**

As with RecQ helicases (Section 1.4.1), Pif1 helicases participate in a wide variety of nucleic acid processes. *S. cerevisiae* Pif1 inhibits telomerase, with *pif1* null mutants possessing longer telomeres and repairing DSBs by aberrantly adding telomeres to the DSB ends (Makovets & Blackburn, 2009; Schulz & Zakian, 1994; Zhou *et al.*, 2000). Several lines of data suggest that the *S. pombe* Pif1 protein and one of the three isoforms of the *S. pombe* Pif1 protein Pfh1p functions in the DNA damage response. Pfh1p levels increase following camptothecin treatment (Pinter *et al.*, 2008), an inducer of DNA lesions through its inhibition of topoisomerase I (Hsiang *et al.*, 1985; Liu *et al.*, 2000). In addition, Pfh1 mutants in *S. pombe* and Pif1 mutants in *S. cerevisiae* are both more sensitive to the DNA damaging agents MMS and hydroxyurea (Tanaka *et al.*, 2002; Wagner *et al.*, 2006). Both Pif1 and Pfh1 localise to DNA damage foci,

Pfh1 after gamma irradiation and Pif1 after treatment with the DNA damaging agent camptothecin (Pinter *et al.*, 2008; Wagner *et al.*, 2006). Furthermore, *S. cerevisiae* Pif1 has been proposed to play a direct role in the repair of DSBs by BIR (Chung *et al.*, 2010; Wilson *et al.*, 2013).

G4 structures are thought to be acted upon by Pif1 helicases. *S. cerevisiae* Pif1 suppresses G4-associated genomic instability (Paeschke *et al.*, 2013) and recombinant *S. cerevisiae* PIF1, along with over 20 other Pif1 helicases, have been shown to bind and/or unwind G4 structures *in vitro* (Paeschke *et al.*, 2013; Ribeyre *et al.*, 2009).

Pif1 helicases play a number of roles in replication. Nuclear *S. cerevisiae* Pif1 participates in the processing of Okazaki fragments and some reports suggest that *S. cerevisiae* Rrm3 is a component of the replisome (Azvolinsky *et al.*, 2006; Matsuda *et al.*, 2007). However, there is debate as to whether Rrm3 is a part of the replisome or if it is specifically recruited to stalled replication forks (Calzada *et al.*, 2005). A major function of *S. cerevisiae* Rrm3 is in the disruption of protein-DNA complexes at certain genomic loci, known as Rrm3-sensitive sites. In the absence of Rrm3, the replication fork pauses at these sites (Ivessa *et al.*, 2002).

Lastly, Pif1 helicases also perform mitochondrial functions. *S. cerevisiae* Pif1 has two isoforms, produced by initiation of translation from two different start sites. Translation initiation from the first start site results in the inclusion of a mitochondrial targeting sequence in the protein and synthesis of the mitochondrial isoform (mtPif1) (Zhou *et al.*, 2002). *S. cerevisiae* mitochondrial Pif1 is required for the maintenance of mitochondrial (mt) DNA, with mitochondrial DNA being lost more quickly in *Pif1* mutant strains (Schulz & Zakian, 1994; Vandyck *et al.*, 1992)

*S. cerevisiae* mtPif1 may be part of the mitochondrial replisome (Cheng *et al.*, 2007; Foury & Vandyck, 1985) and be involved in mitochondrial DNA replication and recombination, though it plays a non-essential role (Cheng *et al.*, 2009).

### 1.4.2.1 Pif1 helicases in *T. brucei*

In contrast to the single Pif1 family helicase possessed by most eukaryotes, *T. brucei* has eight Pif1 helicases, TbPIF1-8. The predicted protein domain organisation of Tb1-8 are shown in Fig. 1-14. The majority of *T. brucei* Pif1 helicases (TbPif1-2 and TbPif4-7) contain a predicted Pif1 family domain, as predicted using InterProScan sequence search. Two helicase domains (though not specifically Pif1 helicase domains) are predicted in TbPif3 and TbPif8 contains a P-loop containing nucleotide triphosphatase hydrolase domain. The absence of a detectable Pif1 helicase domain in two *T. brucei* Pif1 helicases demonstrates the divergent nature of *T. brucei*. The majority of these Pif1 helicases have been shown to play a kinetoplast role: TbPIF1, 2, 5 and 8 have been characterised and are involved in kinetoplast replication (Liu *et al.*, 2009a; Liu *et al.*, 2009b; Liu *et al.*, 2010a; Wang *et al.*, 2012). TbPIF3, 4 and 7 yet to be characterised, but TbPIF3 localises to the cytoplasm while TbPIF4 and 7 localise mainly in the kDNA disc (Liu *et al.*, 2009a). TbPIF6 has been demonstrated to localise to the nucleus (Liu *et al.*, 2009a) but otherwise has not been characterised. All three proteins play roles in genome maintenance, stability or repair in other organisms but have not been characterised in *T. brucei*.

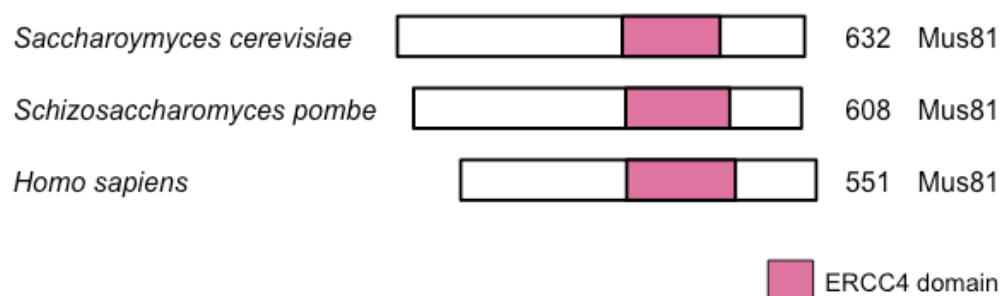
## 1.5 Mus81 endonucleases in DNA repair and a putative *T. brucei* Mus81 homologue

Mus81 (MMS UV Sensitive) is a structure specific endonuclease first identified in *S. cerevisiae* (Boddy *et al.*, 2000; Interthal & Heyer, 2000). It acts in complex with its partner protein (Eme1 in humans and *S. pombe*, Mms4 in *S. cerevisiae*), which is essential for Mus81 nuclease activity (Boddy *et al.*, 2001; Ciccina *et al.*, 2003; Kaliraman *et al.*, 2001; Ogrunc & Sancar, 2003). Mus81 proteins of *S. cerevisiae*, *S. pombe* and *H. sapiens* are shown in Figure 1-15. MUS81 proteins contain an ERCC4 domain. ERCC4 domains are involved in nucleic acid binding and are found in a number of DNA repair proteins, including RAD1 and XPF, as well as in MUS81 and another crossover junction endonuclease, EME1 (Ciccina *et al.*, 2008).

Mus81 acts with its partner protein to perform a number DNA repair functions involving resolution of recombination intermediates arising from replication and recombination by DNA cleavage (Hanada *et al.*, 2007; Regairaz *et al.*, 2011b). Mus81 resolves stalled and collapsed replication intermediates in human and yeast cells (Doe *et al.*, 2002; Hanada *et al.*, 2007; Regairaz *et al.*, 2011a) and promotes the restart of stalled replication forks by fork cleavage (Pepe & West, 2014a). Mus81 also acts on 3' DNA flaps and nicked HJs (Pepe & West, 2014b) and also acts on intact HJs (Boddy *et al.*, 2001; Chen *et al.*, 2001b; Gaillard *et al.*, 2003). Resolution of recombination intermediates by Mus81 results in crossover and yeast *mus81* mutants have a lower rate of meiotic crossover (Boddy *et al.*, 2001; Kaliraman *et al.*, 2001; Osman *et al.*, 2003). Mus81-mediated meiotic crossover has also been observed in mice (Holloway *et al.*, 2008). Additionally, Mus81 acts to promote BIR repair of DSBs by its recruitment to the invading DNA strand, where its endonuclease activity facilitates the establishment of the replication fork required to complete BIR (discussed in Section 1.2.2.2). In this process, Mus81 also promotes template switching, where the invading DNA strand 'switches' to a new homologous template, which frequently results chromosomal translocations (Pardo & Aguilera, 2012; Štafa *et al.*, 2014).

MUS81 is also involved in the response to some forms of DNA damage such as methylation and UV-induced DNA damage in *S. cerevisiae* (Interthal & Heyer, 2000), and interstrand crosslinks in mammalian cells as part of the MUS81-EME1 complex (Hanada *et al.*, 2006). Budding and fission yeast toleration of replication defects also requires Mus81, where Mus81 allows replication through damaged DNA and helps recruit Cds1, a protein that modulates the activity of damage tolerance enzymes (Boddy *et al.*, 2000; Interthal & Heyer, 2000).

The *T. brucei* genome encodes one gene (Tb427.08.6740; TbMUS81) that BLAST (basic local alignment search tool) analysis conducted prior to this work suggested is a Mus81 homologue but that has not been characterised or examined.



**Figure 1-15** Illustration of yeast and human Mus81 proteins. Comparison of the size of eukaryotic Mus81 proteins and location of ERCC4 domains. The number of amino acids in each protein sequence is indicated on the right. Accession numbers can be found in Appendix 7.2. Not to scale.

## 1.6 Aims of this project

The overall aim of this thesis was to investigate the role of several uncharacterised factors in DNA repair and VSG switching in *T. brucei*. In particular, it was aimed to investigate the hypothesis proposed by Boothroyd *et al.* (2009), and since commencing this work supported by Glover *et al.* (2013a), that DSBs at the 70 bp repeats of the active VSG BES initiate VSG switching by recombination. These factors were selected based on the DNA repair and replication functions of their equivalent proteins in other eukaryotes, notably humans and yeast.

The aforementioned uncharacterised *T. brucei* factors were: TbRECQ1 and TbRECQ2 (Tb427.06.3580& Tb427.08.6690; RecQ-like helicases), TbMUS81 (Tb427.08.6740; structure-specific endonuclease) and TbPIF1 (Tb427.10.910; Pif1 family helicase). The specific aims of this work and the approaches taken to address each are detailed below.

### 1.6.1 Roles of four factors in DNA repair

**Aim 1:** Investigate whether each of these four factors are involved in *T. brucei* DNA repair.

**Approach:** Generation of heterozygous and null knockout mutants in wild type cells and analysis of their DNA damage phenotypes.

### **1.6.2 Roles of four factors in DSB repair**

Aim 2: Establish whether each of the four factors are involved in the repair of DSBs, including DSBs at the 70 bp repeats, the location hypothesised to be where VSG switching-initiating DSBs occur.

Approach: Generate heterozygous and null knockout mutants in cell lines that allow the induction of DSBs at specific locations using the I-SceI meganuclease system, and analyse survival and repair of DSBs in these mutants.

### **1.6.3 Roles of four factors in VSG switching**

Aim 3: Assess whether and how the four factors function in VSG switching.

Approach: Generate heterozygous and null knockout mutants in a cell line permitting the VSG switching rate and profile of *T. brucei* cells to be assayed *in vitro* and perform VSG switching experiments using these mutants.

### **1.6.4 DSBs at the 70 bp repeats as the initiators of VSG switching**

These aims sought to investigate the role of these four factors in DSB repair at the site hypothesised to initiate VSG switching (using I-SceI directed DSBs), while also investigating their role in VSG switching in an unbiased fashion (using a VSG switching assay). This concurrent analysis was designed to not only investigate whether and how these factors are involved in DNA repair and VSG switching, but also further our understanding of the initiation of VSG switching and explore the hypothesis that DSBs at the 70 bp repeats are the initiating lesion.

## **Chapter 2**

# **Materials and Methods**



## 2 Materials and Methods

### 2.1 Trypanosome culture

#### 2.1.1 Trypanosome strains and their growth

The trypanosome strain used in this thesis is Lister 427 (McCulloch *et al.*, 1997; Rudenko *et al.*, 1996). This is a monomorphic *T. brucei brucei* strain, usually only displaying long slender blood stream form (BSF) cells. The VSG switching frequency of this strain is  $10^{-6}$ - $10^{-7}$  switches per cell per generation (Lamont *et al.*, 1986; Turner, 1997). For RNAi experiments, the 2T1 strain, derived from the Lister 427 strain, was used (Alsford *et al.*, 2005b).

HMI-9 medium (Hirumi & Hirumi, 1989) was used for *in vitro* growth of BSF Lister 427 trypanosomes. Pre-formulated HMI-9 powder (Gibco, cat. no. 07490915N) was supplemented with 3% w/v NaHCO<sub>3</sub> and 200 µM β-mercaptoethanol to produce 'incomplete' HMI-medium. Incomplete HMI-9 medium was filter sterilized using a 0.22 µm bottle top filter unit (Millipore, cat. no. SCGPS05RE) and stored at 4 °C. For use, HMI-9 was completed with 20% v/v foetal bovine serum (FBS) (Sigma, cat. no. F9665) to produce 'complete' HMI-9. For I-Scel cell lines, tetracycline-free FBS was used (Invitrogen, cat no. 10270-106) and for RNAi lines incomplete HMI-9 was supplemented with 10% v/v tetracycline-free FBS to produce 'complete' HMI-11. After addition of FBS, media was re-filtered before use and stored at 37 °C for no longer than one month. VSG switching analysis experiments were performed using thymidine-free media (see 2.2.4.1). Where appropriate, medium was supplemented with antibiotics at the point of culture preparation at the following concentrations: Hygromycin (Invivogen, cat. no. ant-hg-5b) 5 µg.mL<sup>-1</sup> for I-Scel lines and 1 µg.mL<sup>-1</sup> for RNAi lines, G418 (Invivogen, cat. no. ant-gn-5b) 5 µg.mL<sup>-1</sup>, blasticidin (Invivogen, cat. no. ant-bl-5b) 10 µg.mL<sup>-1</sup>, puromycin (Invivogen, cat. no. ant-pr-5b) 0.2 µg.mL<sup>-1</sup>, phleomycin (Invivogen, cat. no. ant-ph-5b) 2 µg.mL<sup>-1</sup> and 0.5 µg.mL<sup>-1</sup> for RNAi lines. For culture of double knockout cell lines, half of the above concentration of blasticidin and neomycin were used, based on experience of our laboratory.

*In vitro* growth of trypanosomes in HMI-9 or HMI-11 media was carried out at 37 °C in a humidified 5% CO<sub>2</sub> incubator. The population doubling time of the

Lister 427 strain is eight hours (Proudfoot & McCulloch, 2006). Cultures were subcultured regularly to avoid overgrowth and maintained at a maximum of  $2.5 \times 10^6$  cells.mL<sup>-1</sup> in tissue culture plates or vented flasks laid flat. Culture density was monitored using a haemocytometer whereby 10 µL of a culture of parasites under a 1 mm square area was counted and multiplied by  $10^4$  to obtain the number of cells.mL<sup>-1</sup> in the culture.

### **2.1.2 Stabilate preparation and retrieval**

Trypanosome cultures were grown to a density of approx  $1 \times 10^6$  cells.mL<sup>-1</sup> were used to prepare stabilates for long term storage. Two hundred microlitres of a 50:50 solution of HMI-9 or HMI-11 media was added to 800 µL of trypanosome culture in a 1.8 mL cryovial (Alpha Laboratories cat. no. LW3534). Stabilates were wrapped in cotton wool and stored at -80 °C for 24 hours before transfer to liquid nitrogen storage. For retrieval, stabilates were thawed at room temperature and added to nine millilitres of drug-free HMI-9 or HMI-11. Appropriate selective drugs were added the following day and cells were passaged as normal.

### **2.1.3 Transfection of trypanosomes**

Transfection of trypanosomes was performed using the Human T-Cell Nucleofector® Kit (Amaya cat. no. VAPA-1002). Trypanosome cultures were grown to a density of  $\sim 1 \times 10^6$  cells.mL<sup>-1</sup> and  $30 \times 10^6$  were harvested by centrifugation at  $1620 \times g$  for 10 minutes. Cell pellets were resuspended in transfection buffer (supplied with kit), transferred to an electroporation cuvette (supplied with kit) and DNA to be transfected was added. Electroporation was performed using the Amaya Human T Cell Nucleofector® Kit (Amaya, cat. no. VPA-1002) and the electroporated cells added to 30 mL of drug-free HMI-9 medium. Two serial 1:10 dilutions of this cell-medium solution were made and each of the three dilutions was distributed in one millilitre aliquots into 24-well plates. Transfected cells were allowed to recover for 24 hours before being subjected to antibiotic selection.

Selective antibiotic concentrations were as follows: blasticidin 10 µg.mL<sup>-1</sup>, G418 5 µg.mL<sup>-1</sup>, hygromycin 10 µg.mL<sup>-1</sup> unless otherwise stated. For second-

round transfection of knockout constructs into heterozygous lines, the blasticidin and neomycin concentrations were halved, based on our laboratory's experience.

Transformants were counted after 7-10 days by examining the plates under a light microscope and counting the number of wells containing live parasites. Each well containing live cells can be considered clonal if the number of positive wells on that plate is <80% (Wickstead *et al.*, 2003). Six wells were usually expanded for further analysis.

#### **2.1.4 Preparation of trypanosomes for molecular biology procedures**

The required number of trypanosomes (as described below) was harvested by centrifugation at 1620 xg for five minutes at room temperature. Cell pellets were washed once with one millilitre of 1x PBS (phosphate buffered saline) then processed or stored for extraction of genomic DNA, RNA or protein as outlined below.

### **2.2 Analysis of trypanosome phenotypes**

#### **2.2.1 Cell cycle analysis**

For cell cycle analysis, trypanosomes were prepared for microscopy and stained with DAPI (4, 6-diamidino-2-phenylindole) (SouthernBiotech, cat no 0100-20) (section 2.11.1). Differential interface contrast (DIC) was used to visualise whole cells and UV to visualise DAPI. Images were acquired using Volocity (Perkin Elmer) and images analysed using ImageJ. Cells were counted according to the number of nuclei and kinetoplasts they contained. Cells in the G1 phase of the cell cycle contain one nucleus and one kinetoplast (1N1K). Kinetoplast division precedes nuclear division, thus cells in the S phase of the cell cycle contain one nucleus and two kinetoplasts (1N2K). After nuclear division, G2 phase cells contain two nuclei and two kinetoplasts (2N2K). Any cells not conforming to 1N1K, 1N2K or 2N2K were classified as 'other'.

## 2.2.2 Analysis of trypanosome *in vitro* growth

*In vitro* growth analysis was performed by inoculating 5 mL of medium with cells, to a final density of  $1 \times 10^4$  cells.mL<sup>-1</sup> or  $5 \times 10^4$  cells.mL<sup>-1</sup>. The inoculating cultures were previously grown to a maximum of  $1 \times 10^6$  cells.mL<sup>-1</sup>. Cell density was counted at 24, 48, 72 and 96 hours. Growth analysis was repeated at least three times for each cell line and results plotted on a semi-logarithmic scale.

## 2.2.3 Analysis of DNA damage sensitivity

Analysis of DNA damage sensitivity was performed using three DNA damaging agents: methyl methanesulfonate (MMS) (Sigma cat. no. M4016), phleomycin (Invivogen cat. no. ant-ph-5b) and hydroxyurea (Sigma cat. no. H8627). DNA damaging agents were used as described for each procedure and in all cases were added to the culture for chronic exposure and not washed off.

MMS is an alkylating agent that methylates purines at the 7' position of guanosine residues and the 3' position of adenine residues (Brookes & Lawley, 1961; Reiter *et al.*, 1972). Phleomycin is an antibiotic that blocks the activity of DNA polymerase, inhibiting DNA synthesis, resulting in formation of single and double strand DNA breaks (SSBs and DSBs, respectively), though primarily double strand breaks (Falaschi & Kornberg, 1964; Reiter *et al.*, 1972). Hydroxyurea inhibits ribonucleotide reductase, an enzyme involved in dNTP synthesis. This results in a depletion of the cellular dNTP pool and causes stalling of DNA replication (Bianchi *et al.*, 1986).

### 2.2.3.1 DNA damage growth curve analysis

Growth curves were set up from cultures grown to a density not exceeding  $1 \times 10^6$  cells.mL<sup>-1</sup>. The day prior to set up of the growth curve, all cultures were passaged into medium without selective antibiotics. Assays were set up and analysed as in Section 2.2.2 with the modification that after preparation of starting cultures, the culture was aliquoted and the appropriate DNA damaging agent added to the appropriate final concentration as follows. MMS: untreated, 0.0002%, 0.0003% and 0.0004%. Phleomycin: 0.05 µg.mL<sup>-1</sup>, 0.075 µg.mL<sup>-1</sup> and 0.1 µg.mL<sup>-1</sup>. Hydroxyurea: 0.01 mM, 0.02 mM and 0.03 mM. Culture density was

counted at 24, 48 and 72 hours and DNA damaging agent added again to the same final concentration if cultures required dilution (where specified).

### **2.2.3.2 DNA damage clonal survival assays**

The day prior to set up of the clonal survival assays all cultures were passaged into drug-free medium. When setting up the assay, cultures were diluted to 1 cell/200  $\mu\text{L}$  ( $0.5 \times 10^1 \text{ cells.mL}^{-1}$ ) and distributed in 200  $\mu\text{L}$  aliquots among three 96 well culture plates after adding phleomycin, hydroxyurea or MMS at the appropriate concentration, or left untreated. Concentrations of DNA damaging agents used were: MMS 0.0001%, 0.0002%, 0.0003% and 0.0004%; hydroxyurea: 0.01 mM, 0.02 mM and 0.03 mM; phleomycin: 0.1  $\mu\text{g.mL}^{-1}$ , 0.2  $\mu\text{g.mL}^{-1}$  and 0.3  $\mu\text{g.mL}^{-1}$  or 0.05  $\mu\text{g.mL}^{-1}$ , 0.10  $\mu\text{g.mL}^{-1}$  and 0.15  $\mu\text{g.mL}^{-1}$ .

After ~10 days incubation, or when cultures were saturated, the number of surviving clones was counted on each plate and the mean survival (of the three 96 well plates) for each damaging agent concentration was calculated. Positive wells (with live cells) were determined by the colour of the medium: HMI medium turns yellow at high cell density. This was graphed as percentage survival relative to untreated cells. The mean of the relative percentage survival was calculated from multiple experiments and used to present combined data from multiple experiments.

### **2.2.3.3 I-SceI double strand break survival assays**

Assays were carried out as in (Glover *et al.*, 2013a). Cultures were diluted at sub-clonal dilution (HR1 cell lines:  $0.15 \times 10^1 \text{ cells.mL}^{-1}$ , HRES cell lines:  $0.26 \times 10^1 \text{ cells.mL}^{-1}$ ), divided into two aliquots then tetracycline (Calbiochem cat. no. 583411) added to one (final concentration 2  $\mu\text{g.mL}^{-1}$ ) to induce I-SceI expression. Cultures were distributed in 200  $\mu\text{L}$  aliquots into 96 well plates (four plates each of uninduced and induced cells). After 7-10 days incubation (before cultures were saturated) the number of surviving clones was counted and survival was normalised to un-induced cultures. Positive wells (with live cells) were determined by the colour of the medium: HMI medium turns yellow at high cell density.

From each uninduced cell line, six surviving clones were expanded for subsequent analysis. Twenty-five surviving induced HR1 and all surviving induced HRES clones were expanded. Selected clones were tested for puromycin sensitivity (final concentration  $1 \mu\text{g} \cdot \text{mL}^{-1}$ ) and genomic DNA was extracted for PCR analysis. HR1 cell lines were assayed using PCR for MMEJ by amplification across the *RFP:PURO* cassette and HRES cell lines were assayed by PCR for *VSG221* (primers #90 and #91, 955 bp product) and *ESAG1* (primers #92 and #93, 328 bp product) loss. Primer sequences are listed in Appendix 7.1. Genomic DNA preparations were also assayed to check the presence of amplifiable genomic DNA by amplification of a gene not knocked out in that cell line. This positive control gene was different depending on the knockout cell line and is specified (along with PCR primers used) in each experiment.

## 2.2.4 VSG switching analysis

### 2.2.4.1 Preparation of media for the culture of VSG switching cell *GFP221hygTK*

The bloodstream form *T. brucei* cell line generated for analysis of VSG switching (*GFP221hygTK*) was cultured in thymidine-free medium. This consisted of Isocove's Modified Dulbecco's Medium (Invitrogen, cat no. 21980-065) supplemented with 20% FBS (Sigma, cat. no. F9665), 1 mM hypoxanthine (Sigma, cat no. H9377), 0.05 mM bathocuproine disulphonic acid (Sigma, cat no. B1125), 1 mM sodium pyruvate (Sigma, cat no. P5280), 1.5 mM L-cysteine (Sigma-Aldrich, cat no. C7352) and 200  $\mu\text{M}$   $\beta$ -mercaptoethanol. Media was filter sterilised prior to use.

### 2.2.4.2 Analysis of VSG switching frequency and mechanism

Mutants were generated in the *GFP221hygTK* cell line to analyse VSG switching frequency and mechanism. This cell line contains two constructs integrated into the active *VSG* BES (MiTat 1.2); *eGFP* and puromycin are integrated downstream of the promoter and a hygromycin resistance gene fused to the thymidine kinase (*TK*) gene is integrated between the 70 bp array and the active *VSG*. Negative selection against thymidine kinase using ganciclovir allows selection of cells that no longer express the *TK* gene, either through a VSG switching event or thymidine kinase point mutation.

Cells were passaged to a density of  $1 \times 10^4$  cells.mL<sup>-1</sup> either in the presence or absence of puromycin ( $0.2 \mu\text{g.mL}^{-1}$ ) and incubated for 48 hours to allow them to switch VSG. After 48 hours cultures were diluted to  $1.25 \times 10^4$  cells.mL<sup>-1</sup> in the presence of  $4 \mu\text{g.mL}^{-1}$  ganciclovir (Sigma-Aldrich, cat. no. G2536) and plated in 200  $\mu\text{L}$  aliquots over three 96 well plates, resulting in  $2.5 \times 10^3$  cells per well. After one week, surviving clones (in all experiments apart from the first RECQ2 VSG switching experiment: 20 each from wild type and heterozygote (+/-) lines and 30 from double knockout (-/-) cell lines) were selected on a random basis to expand in thymidine- and drug-free medium for further analysis. Plates were incubated for a further week before the number of surviving clones on each plate was counted. Cells were collected for genomic DNA extraction for PCR analysis and for SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) for western blot analysis.

To calculate the VSG switching frequency the number of surviving clones in total (all three 96 well plates) was divided by the total number of cells plated to obtain the number of switching events per cell. This number was then divided by the number of generations in the 48 hour incubation prior to plating the cells to obtain the VSG switching rate per cell per generation. The number of generations was calculated for each cell line in the presence and absence of puromycin using the cell density measured at the 48 hour time point. To determine the number of generations, the number of cell doublings per hour was calculated. The doubling time was then calculated and was then used to calculate the number of generations during the 48 hour incubation. These formulae are shown in Figure 2-1. In the formula in Fig. 2-1 the number 0.301 is used because this is the log of two and division occurs by binary fission, therefore each division results in a doubling of the population size.

$$\text{Cell doubling per hour (k)} = \frac{(\text{Log}_{10} \text{ cells.mL}^{-1} \text{ at 48 hours}) - (\text{Log}_{10} \text{ cells.mL}^{-1} \text{ at 0 hours})}{0.301 \times 48 \text{ hours}}$$

$$\text{Doubling time (g)} = \frac{1 \text{ (hour)}}{k}$$

$$\text{Number of generations (n)} = \frac{48 \text{ hours}}{g}$$

**Figure 2-1**      **Formulae used to calculate the number of generations in VSG switching experiments**

These formulae were used to calculate the number of cell doublings per hour, time taken to double the population and the number of generations during the 48 hour VSG switching period. The number of generations was then used to calculate the VSG switching frequency, as described in the text.

SDS-PAGE gels were loaded with  $2.5 \times 10^6$  cells per lane. SDS-PAGE was performed as outlined in Section 2.5.2 and western blotting was performed as outlined in Section 1.6. Blots were probed with rabbit anti-VSG221 antisera (Gift, David Horn, (Glover *et al.*, 2013a)) at a 1:20,000 dilution and rabbit anti-GFP antisera (Abcam, cat. no. ab290) at a 1:5000 dilution. Goat anti-rabbit HRP conjugated antisera (Molecular Probes, cat. no. G-21234) was used as secondary antisera. Genomic DNA was assayed for the presence of *VSG221* and *GFP* by PCR using primer pairs #156 & #157 (522 bp product) and #160 & #161 (361 bp product), respectively. As a positive control to test the presence of amplifiable genomic DNA, a 453 bp region of RNA polymerase I was amplified using primer pairs #158 & #159.

## 2.3 Trypanosome RNA interference (RNAi)

### 2.3.1 Cloning of *RECQ1* RNAi plasmid using Gateway® system

The modified pRPa<sup>ISL</sup> plasmid, pGL2084, developed by Jones *et al.* (2014) was used to clone plasmids for RNAi. The pGL2084 plasmid contains two sets of inverted Gateway donor sites (containing attP sites). Each pair of Gateway donor sites flank a *ccdB* counter selectable marker. The two sets of Gateway donor sites are separated by 150 bp of the *lacZ* gene. The *RECQ1* RNAi target sequence was determined using the TrypanoFAN: *RNAi*t programme (Redmond, 2003). Primers #67 & #68 incorporated attB1 and attB2 sites to facilitate recombination with the pGL2084 attP sites. The resultant PCR product was cloned into pGL2084 in a BP recombinase reaction (Invitrogen, cat. no. 11789013) and propagated using Max Efficiency DH5α Competent Cells (Life Technologies, cat. no. 18258012). Prior to transformation into 2T1 cells (Alsford *et al.*, 2005a), the *RECQ1* RNAi plasmid was linearised by digestion with the restriction enzyme PacI.



## **2.4 Isolation of material from trypanosomes**

### **2.4.1 Isolation of genomic DNA**

Cell pellets obtained as described above (approximately  $3 \times 10^6$  cells per sample) were frozen at  $-20^{\circ}\text{C}$  prior to genomic DNA extraction. DNA was extracted using the Qiagen DNeasy Blood and Tissue kit (Qiagen cat. No. 69606) following the manufacturer's instructions. DNA concentration was determined photospectrometrically using a Nanodrop 1000 (Thermo Scientific).

### **2.4.2 Isolation of total RNA**

Cell pellets obtained as above (approximately  $2 \times 10^7$  cells per sample) were frozen at  $-80^{\circ}\text{C}$  prior to RNA extraction. RNA was extracted using the Qiagen RNeasy kit (Qiagen cat. no. 74106). DNase digestion was performed after RNA extraction using the Qiagen RNase-Free DNase Set (Qiagen cat. no. 79254) and RNA clean up performed afterwards using the Qiagen RNeasy kit, following the manufacturer's instructions. Total RNA preparations were stored at  $-80^{\circ}\text{C}$ .

## **2.5 Electrophoresis**

### **2.5.1 DNA electrophoresis**

Separation of DNA fragments for routine analysis was achieved by loading on 1.2% (w/v) agarose gels (Invitrogen, Paisley, UK, cat. No. 15510-027) run at 100V in 1 x TAE buffer (40 mM Tris, 19 mM acetic acid, 1 mM EDTA). SYBR-Safe DNA Stain (Life Technologies, cat. no. S33102) was used at a 1:20,000 dilution to visualise DNA under ultraviolet light. Gels were imaged using Bio-Rad GelDoc (BioRad) and images captured using Quantity One (Bio-Rad).

### **2.5.2 Protein electrophoresis**

SDS-PAGE was used for separation of proteins in crude cell lysates using pre-cast NuPAGE® Novex® 3-8% Tris-acetate gels (Life Technologies, cat. no. EA0375BOX). Trypanosome cell pellets were frozen at  $-80^{\circ}\text{C}$  until ready to be electrophoresed. Cell pellets were resuspended in protein loading buffer (1x Life Technologies NuPAGE® LDS sample buffer cat. no. NP0007, 2.5%  $\beta$ -mercaptoethanol, 2x Roche cOmplete Mini protease inhibitor cocktail tablets

cat. no. 04693124001), mixed well and boiled at 100°C for five minutes. Approximately  $5 \times 10^6$  cell equivalents were loaded per lane and gels were run at 100V for 50 minutes. For estimation of protein molecular weight, 10 µL of HiMark™ Pre-stained Protein Standard (Life Technologies, cat. no LC5699) was also loaded.

## **2.6 Blotting**

### **2.6.1 Western blot transfer**

Western blotting of SDS-PAGE gels was carried out using the Mini Trans-Blot® Cell (Bio-Rad). Gels, Hybond ECL nitrocellulose membrane (Amersham, cat. no. RPN303D), foam and filter paper (Bio-Rad) were equilibrated in transfer buffer (25 mM Tris pH 8.3, 192 mM Glycine, and 20% (v/v) methanol)) prior to assembling the gel sandwich. The gel sandwich comprised the gel and nitrocellulose membrane, surrounded by filter paper and foam, sandwiched into a plastic cassette. An ice block was placed next to the cassette to prevent overheating. The transfer gel sandwich was electrophoresed at 400 mA for one hour followed by transfer at 90 mA overnight at 4°C. The success of transfer was determined by incubation of the membrane in ponceau solution (Sigma-Aldrich, cat. no. P7170) for 10 minutes and visualised using Bio-Rad GelDoc (BioRad) and images captured using Quantity One (Bio-Rad).

### **2.6.2 Western blot detection**

### **2.6.3 Hybridisation and detection of antibodies**

Western blot membranes were washed once in 1x PBS then incubated in blocking solution (1x PBS, 5% Milk (Marvel), 0.1% Tween-20 (Sigma-Aldrich cat. no. P1379)) for one hour or overnight at 4°C, on a rocker. Membranes were then placed in blocking buffer containing primary antisera (Table 2-1) for one hour. Membranes were rinsed twice in PBST (PBS, 0.1% Tween-20 (Sigma)) for 10 minutes, then placed in blocking buffer containing the secondary antisera (Table 2-1) for one hour. Secondary antisera were horseradish peroxidase conjugated. Membranes were washed three times in PBST for 10 minutes each before applying SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, cat. no. 34078). Membranes were incubated in the dark for one

minute after application of the substrate, then exposed to an X-ray film (Kodak) for 10 minutes to overnight. X-ray films were visualized by developing in a Kodak M-25-M X-omat processor.

#### **2.6.4 Stripping western blots**

Nitrocellulose membranes were stripped by incubating in Restore Western Blot Stripping Buffer (Thermo Scientific, cat. no. 21059) on a rocker for 10-20 minutes. Membranes were washed in PBST before re-probing.

| Experiment                  |            | Antisera name               | Host animal | Target animal | Company/Laboratory | Catalogue no. | Dilution |
|-----------------------------|------------|-----------------------------|-------------|---------------|--------------------|---------------|----------|
| Myc detection (WB)          | Primary    | AntiMyc                     | Mouse       | -             | Millipore          | 05-724        | 1:7000   |
|                             | Secondary  | Anti-mouse HRP conjugated   | Goat        | Mouse         | Life Technologies  | G21040        | 1:5000   |
| Myc detection (IFA)         | Conjugated | Anti-Myc Alexa Fluor 488    | Mouse       | -             | Millipore          | 16-224        | 1:2000   |
| RAD51 detection (IFA)       | Primary    | Anti-RAD51                  | Rabbit      | -             | McCulloch          | -             | 1:1000   |
|                             | Secondary  | Anti-rabbit Alexa Fluor 594 | Goat        | Rabbit        | Life Technologies  | A-11005       | 1:2000   |
| vH2A detection (IFA)        | Primary    | Anti-vH2A                   | Rabbit      | -             | McCulloch/Mottram  | -             | 1:1000   |
|                             | Secondary  | Anti-rabbit Alexa Fluor 594 | Goat        | Rabbit        | Life Technologies  | A-11005       | 1:2000   |
| VSG221 detection (IFA)      | Primary    | Anti-VSG221                 | Rabbit      | -             | Matthews           | -             | 1:1000   |
|                             | Secondary  | Anti-rabbit Alexa Fluor 594 | Goat        | Rabbit        | Life Technologies  | A-11005       | 1:2000   |
| EF1 $\alpha$ detection (WB) | Primary    | Anti-EF1 $\alpha$           | Mouse       | -             | Millipore          | 05-235        | 1:25000  |
|                             | Secondary  | Anti-mouse HRP conjugated   | Goat        | Mouse         | Life Technologies  | G21040        | 1:5000   |
| VSG221 detection (WB)       | Primary    | Anti-VSG221                 | Rabbit      | -             | Horn               | -             | 1:25000  |
|                             | Secondary  | Anti-rabbit HRP conjugated  | Goat        | Rabbit        | Life Technologies  | G21040        | 1:5000   |
| GFP detection (WB)          | Primary    | Anti-GFP                    | Rabbit      | -             | abcam              | ab290         | 1:5000   |
|                             | Secondary  | Anti-rabbit HRP conjugated  | Goat        | Rabbit        | Life Technologies  | G21040        | 1:5000   |
| VSG221 detection (IFA)      | Primary    | Anti-VSG221                 | Rabbit      | -             | Horn               | -             | 1:1000   |
|                             | Secondary  | Anti-rabbit Alexa Fluor 594 | Goat        | Rabbit        | Life Technologies  | A-11005       | 1:2000   |

Table 2-1 List of antisera used in this thesis

The table lists the antisera utilised in this work and the dilutions at which they were used. Horn, David Horn; Matthews, Keith Matthews; McCulloch, Richard McCulloch; Mottram, Jeremy Mottram. WB, western blot; IFA, immunofluorescence.

## **2.7 Molecular biology techniques**

### **2.7.1 Polymerase chain reaction (PCR)**

For diagnostic PCR reactions Taq DNA polymerase (New England Biolabs (NEB) cat. no. M0273S) was used. A typical 20  $\mu$ L reaction contained 0.1  $\mu$ M 5' primer, 0.1  $\mu$ M 3' primer, 1x ThermoPol Buffer (NEB, cat no M0273S), 0.2 mM dNTPs (NEB cat. no. N0446S), 0.2 units Taq polymerase and dH<sub>2</sub>O made up to 25  $\mu$ L. PCRs were performed with typical conditions of 95°C for five minutes, followed by 30-35 cycles of 95°C for one minute, 50-60°C for one minute, 72°C for one minute per 1 Kb of expected product and a final cycle of 72°C for five to ten minutes.

Phusion DNA polymerase (NEB cat. no. M0530S) was used for PCRs where a low error rate was required. A typical 50  $\mu$ L reaction contained 0.1  $\mu$ M 5' primer, 0.1  $\mu$ M 3' primer, 1x Phusion High fidelity (HF) buffer (supplied with Phusion polymerase), 0.2  $\mu$ M dNTPs (NEB cat. no. N0446S), 1 unit of Phusion DNA polymerase and dH<sub>2</sub>O to 50  $\mu$ L. PCRs were performed with typical conditions similar to those for Taq polymerase above, with the modification that the denaturation temperature was 98°C.

### **2.7.2 DNA fragment gel extraction**

When separated DNA was to be extracted from an agarose gel, a clean scalpel was used to excise the DNA fragment, which was then placed in a sterile microfuge tube. DNA was extracted from gel fragments using the Qiagen Gel Extraction Kit (Qiagen cat. no. 28706) following the manufacturer's instructions.

### **2.7.3 cDNA synthesis**

RNA was used as the template for cDNA synthesis. cDNA synthesis was performed using the Primer Design Precision nanoScript Reverse Transcription kit (Primer Design cat. no. RT-nanoScript) according to manufacturer's instructions. Random primers (supplied by manufacturer) were used for cDNA synthesis. A negative control, to which no reverse transcriptase enzyme was added, was also performed in order to determine whether DNase digestion performed during RNA extraction was complete.

### 2.7.4 Restriction enzyme analysis of DNA

All restriction enzyme digestions were performed using enzymes supplied by New England Biolabs in buffers recommended by the manufacturer. In a typical digestion, 10 units of enzyme were added per microgram of DNA and incubated at 37°C for one hour, or overnight in the case of digesting plasmids, prior to trypanosome transformation.

### 2.7.5 Primer design

*T. brucei* gene sequences were downloaded from TriTrypDB ([www.tritrypdb.org](http://www.tritrypdb.org)) using the gene accession number. Sequences were viewed in CLC Genomics Workbench 6 (CLC Bio). Primers were designed using CLC Genomics Workbench 6 and the web-based programme Primer3 (<http://frodo.wi.mit.edu/primer3/>). Restriction sites were added to the 5' end of primers with additional bases at the 5' terminus for efficient restriction enzyme digest of PCR products.

Start and stop codons were added or removed as necessary and it was verified that the resulting products would be in-frame with fusion tags. All oligonucleotides in this thesis were synthesised by Eurofins MWG Operon ([www.eurofinsdna.com](http://www.eurofinsdna.com)) and are listed in Appendix 7.1, Table 7-1.

### 2.7.6 Plasmid sequencing

All sequencing was performed by MWG Eurofins Operon (Ebersburg, Germany).

### 2.7.7 Concentration of digested DNA

Plasmids digested prior to trypanosome transfection were concentrated using the Zymogen DNA Clean and Concentrator Kit (Zymogen cat. no. D4005 & D4003) following the manufacturer's instructions.

## 2.8 Cloning

### 2.8.1 T4 DNA ligase

DNA ligation was used to insert digested PCR products into plasmids. A typical 20 µL reaction contained: 2.5 units of T4- DNA ligase (NEB, cat. no. M0202),

molar ratios of vector to insert was typically 1:3 (typically 50 ng linearised vector: 150 ng digested insert DNA), 2  $\mu$ L of 10x ligation buffer (supplied with T4 DNA ligase) and dH<sub>2</sub>O to a final volume of 20  $\mu$ L. The ligation mix was incubated at room temperature overnight.

## 2.9 Transformation of *E. coli* and plasmid retrieval

### 2.9.1 Transformation of *E. coli*

For general transformations, chemically competent DH5- $\alpha$  cells were used. Cells stored at -80°C were thawed on ice and 5-10  $\mu$ L of ligation reaction or 50-300 ng of purified plasmid was added to 50  $\mu$ L of cells in a sterile microfuge tube. Transformations were incubated on ice for 30 minutes before being heat shocked at 42°C for 30 seconds. The cells were left to recover on ice for five minutes before 1 mL of SOC medium (5 g yeast extract, 20 g tryptone, 0.5 g NaCl, 10 mL 1M MgCl<sub>2</sub>, 10 mL 2M glucose, 10 mL 1M MgSO<sub>4</sub> in 1 L) added and the microfuge tube was incubated at 37°C for one hour. Recovered cells were spread on Luria-Bertani broth (LB; 5 g yeast extract (Formedium), 10 g tryptone (Formedium), 10 g NaCl in 1 L dH<sub>2</sub>O pH 7.0 (NaOH)] and 20 g agar (Formedium)) agar plates containing 100  $\mu$ g.mL<sup>-1</sup> ampicillin (Sigma-Aldrich cat. no. A9518) and incubated overnight at 37°C.

For transformation of plasmids (*221GP1* and *HYG-TK*) used to construct the VSG switching lines *GFP221hygTK* XL-10 Gold Supercompetent Cells (Agilent Technologies cat. no. 200314) were used. Transformations were performed following the manufacturer's instruction and cultured using ZYM-5 agar and medium (see Section 2.9.1.1). Transformation of BP plasmids (for constructing the RNAi cell line) were performed into MAX Efficiency® DH5 $\alpha$  Competent Cells (Life Technologies cat. no. 18258-012) following the manufacturer's instructions.

#### 2.9.1.1 Preparation of ZYM-5 medium

ZYM medium and agar was used to propagate plasmids used to generate the VSG switching cell line *GFP221hygTK*. ZYM-5 medium: 10 g.L<sup>-1</sup> trypton, 5 g.L<sup>-1</sup> yeast extract, 5 g.L<sup>-1</sup> glycerol. Medium was autoclaved and, prior to use, the appropriate volume of autoclaved 50x "stock M" and 500x "stock Mg" was

added. 50x “stock M”: 1.25 M Na<sub>2</sub>HPO<sub>4</sub>, 1.25 M KH<sub>2</sub>PO<sub>4</sub>, 2.5 M NH<sub>4</sub>Cl. 500x “stock Mg”: 1M MgSO<sub>4</sub>.

## 2.9.2 Extraction of plasmid DNA from *E. coli*

Colonies containing the desired plasmid were inoculated into overnight 5 mL cultures of Luria Bertani broth. Approximately 18 hours later, cultures were centrifuged at 2000x g for 10 minutes then processed using the QIAprep Spin Miniprep Kit (Qiagen cat. no. 27104) following the manufacturer’s instructions. DNA concentration was determined photospectrometrically using a Nanodrop 1000 (Thermo Scientific).

## 2.10 Microscopy

### 2.10.1 DAPI staining

Cells were harvested by centrifugation at 405 x g for 10 minutes and the cell pellet washed once with 1x vPBS (NaCl 8 g/L, KCl 2.2 g/L, Na<sub>2</sub>HPO<sub>4</sub> 22.7 g/L, KH<sub>2</sub>PO<sub>4</sub> 4.4 g/L, 15.7 g/L sucrose, 1.8 g/L glucose, pH 7.4). The pellet was resuspended and 20 µL (containing 2 x 10<sup>6</sup> cells) was pipetted onto one well of a 12-well glass slide (Menzel-Gläser, cat. no. X2XER202W) pre-treated with poly-L-lysine (Sigma, cat. no. P8920) and a mini PAP pen (Life Technologies, cat no. 00-8877) was used to create a barrier between wells. Cells were allowed to settle for five minutes, most of the liquid removed with a pipette and 20 µL of 3.7% paraformaldehyde (PFA) added to fix cells. PFA was removed after four minutes. Cells were permeabilised by adding 25 µL Triton X-100 (Thermo Scientific, cat. no. 85112) in PBS for 10 minutes. Fifty microlitres of 100 mM glycine was added and incubated for five minutes and repeated to neutralise free aldehyde groups. Wells were washed twice with 50 µL 1x PBS for five minutes. Five microlitres of DAPI (SouthernBiotech, cat no. 0100-20) was added to every well and incubated for three minutes at room temperature before addition of a coverslip and sealing with nail varnish. Slides were stored at 4°C in the dark. All solutions used were filter sterilised.



## **2.10.2 Immunofluorescence microscopy**

Cells were prepared as in Section 2.11.1, modified thus: Instead of adding DAPI, 25  $\mu$ L of blocking solution (1% BSA (Sigma, cat. no. A2153-50G), 0.2% Tween-20 in PBS) was added to the well and incubated for one hour in a moist chamber. Blocking solution was removed, 25  $\mu$ L primary antisera in blocking buffer added and incubated for one hour in a moist chamber. Primary antisera was removed, 25  $\mu$ L of secondary antisera in blocking buffer added and incubated as before. The well was washed twice with 50  $\mu$ L PBS and DAPI added and slides sealed as in Section 2.11.1. Antibodies used and their dilutions are listed in Table 2-1. Slides were stored at 4°C in the dark. All solutions were filter sterilised.

DIC was used to visualise whole cells and UV was used to visualise DAPI. A fluorescein isothiocyanate (FITC) (520 nm) filter was used to visualise Alexa Fluor 488 conjugated antibodies and a rhodamine (673 nm) filter was used to visualise Alexa Fluor 594 conjugated antibodies. Slides were imaged using a Zeiss Axioskop microscope and images acquired using the Volocity software package (Elmer Perkin). Acquired images were analysed using ImageJ image processing software.

## **2.11 Bioinformatics**

Kinetoplastid sequences were downloaded from TripTrypDB (TripTrypDB.org) and all other gene and protein sequences were downloaded from NCBI. All sequences were viewed in CLC Genomics Workbench 6. Protein BLAST (P-BLAST) analysis of kinetoplastid genomes was carried out on TripTrypDB.org and P-BLAST analysis of all other organisms was performed on NCBI. P-BLAST is a protein sequence similarity search method that compares the input (query) protein sequence to protein sequences in the chosen database (non-redundant protein database based here). The result is the identification of regions of local alignment between the query and database sequences. Results include the aligned sequence, alignment score and length of the aligned region. PSI-BLAST is a protein sequence profile search that utilises the alignment generated in an initial BLAST search. PSI-BLAST then generates a multiple sequence alignment based on the results of the BLAST search and calculates a position-specific score matrix (PSSM). A PSSM describes the sequence conservation of the multiple

sequence alignment, giving higher scores to more highly conserved residues. The P-BLAST search is then repeated using the PSSM as the input sequence and the results of this search are used to refine the PSSM. This process is repeated until no new sequences are detected, known as convergence. Iterative searching using a PSSM enables PSI-BLAST to detect more distantly related sequences than P-BLAST and is therefore beneficial when the input sequence is highly divergent and P-BLAST is insufficient to detect similar proteins. Protein domain prediction was performed using InterProScan sequence search (<http://www.ebi.ac.uk/interpro/search/sequence-search>). InterProScan brings together a number of other databases that contain information about known proteins' function, including the protein families, domains and functional sites. This permits predictive functional characterisation of the user's input sequence.

ClustalX and BoxShade were used to construct multiple sequence alignments (except for alignments used to produce trees). Clustal works by calculating all pairwise alignments and recording the score for each pair of sequences in a distance matrix. A tree using the neighbour joining method is then constructed using this distance matrix, identifying the two most closely related sequences. These two sequences are aligned and the consensus is calculated. The next two most closely related sequences are added to the alignment, the consensus is calculated again and the process is repeated until all sequences are aligned. Construction of trees was performed using CLC Genomics Workbench 6. A multiple sequence alignment was constructed using the entire amino acid sequence of each protein and a tree generated using the neighbour-joining (NJ) method. This tree construction method was chosen as it does not assume a molecular clock and is appropriate when the evolutionary relationship between all sequences is unknown. For this reason, the NJ method was chosen over the unweighted pair group with arithmetic mean (UPGMA) method, which assumes a randomised molecular clock. The NJ method constructs a tree by finding pairs of sequences (neighbours) that result in the minimum total branch length at each stage of the tree construction, therefore producing a tree with the minimum evolutionary distance. It is therefore a 'bottom-up' clustering method.

## **Chapter 3**

### **Analysis of the roles in general DNA repair of RECQ2, PIF6 and MUS81**

## 3 Analysis of the roles in general DNA repair of RECQ2, PIF6 and MUS81

### 3.1 Introduction

Many proteins with functions in DNA break repair have been identified in *T. brucei*, notably RAD51 and RAD51 paralogues, BRCA2 and MRE11. However, only some of the proteins involved in DNA repair been demonstrated to function in VSG switching. These include RAD51 (McCulloch & Barry, 1999), the RAD51-3 paralogue (Dobson *et al.*, 2011; Proudfoot & McCulloch, 2005), BRCA2 (Hartley & McCulloch, 2008), the topoisomerase TOPO3 $\alpha$  (Kim & Cross, 2010) and RMI1 (Kim & Cross, 2011).

This chapter provides evidence for repair functions for three further proteins, which were identified in the *T. brucei* genome through P-BLAST searches using eukaryotic homologue protein sequences: RECQ2, one of two RecQ-like helicases in the *T. brucei*, MUS81, a structure specific endonuclease and PIF6, one of eight Pif1-like helicases encoded by *T. brucei*.

The aim of this chapter was to analyse *RECQ2*, *MUS81* and *PIF6* to determine whether they act in DNA damage repair in BSF *T. brucei*, before continuing in the next chapter to analyse specifically their role in DSB repair and VSG switching.

### 3.2 Sequence analysis of putative DNA repair factors

TbRECQ2, TbPIF6 and TbMUS81 were identified and selected for analysis through P-BLAST searches using eukaryotic homologues of each protein, searching against the *T. brucei* Lister 427 proteome on tritrypdb.org. However, as *T. brucei* is a highly diverged eukaryote, further sequence analysis was undertaken to confirm the identity of the *T. brucei* proteins and investigate their similarity to characterised eukaryotic homologues.

#### 3.2.1 *T. brucei* RECQ2 is a RecQ helicase homologue

To identify TbRECQ1, NCBI protein-BLAST (P-BLAST) searches (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>; NCBI nr November 2014) were

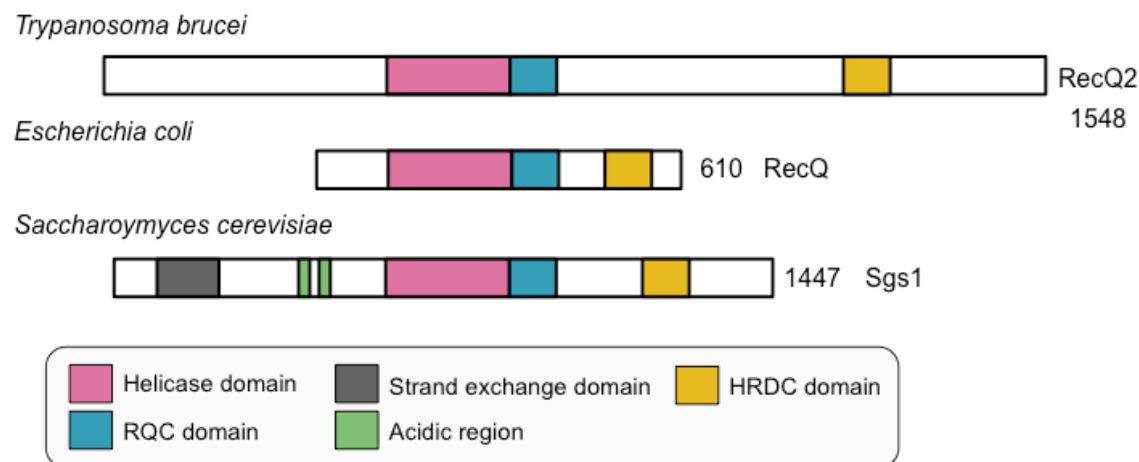
performed (data not shown), using 18 different RecQ protein sequences from *Homo sapiens*, *Mus musculus*, *Arabidopsis thaliana*, *Caenorhabditis elegans* and *Saccharomyces cerevisiae* as the query sequence and using only the *T. brucei* Lister 427 strain as the target organism (<http://tritrypdb.org>; Version 8.1). TbRECQ2 (TriTrypDB accession number Tb427.08.6690) and TbRECQ1 (TriTrypDB accession number Tb427.06.3580) were consistently the only protein sequences retrieved. TbRECQ1 is discussed in Chapter 5. To begin to test the putative function and orthology of TbRECQ2, the predicted protein sequence was used for P-BLAST analysis.

P-BLAST searches using the *T. brucei* RECQ2 amino acid sequence as the query sequence were also performed to ensure that searches returned RecQ helicases. The hits with the highest scores and lowest E values were all RecQ helicases. The top 11 hits against RECQ2 were *Trypanosoma spp* and *Leishmania spp* RecQ-like proteins, with an expectancy value of 0. A number of plant and mammalian RecQ-like proteins were retrieved as good alignments for RECQ2, with E values in the range of  $10^{-120}$  to  $10^{-110}$ . These were primarily mammalian BLM homologues, in addition to several RECQ4A and RECQ4B homologues.

#### Protein domain analysis using InterProScan sequence search

(<http://www.ebi.ac.uk/interpro/search/sequence-search/> (Hunter *et al.*, 2012)) was used to predict the domains present in RECQ2. RECQ2 contains a DEAD/DEAH box helicase domain (approximately aa460-666), involved in ATP and nucleic acid binding, and a C-terminal helicase domain (approximately aa674-825) (RQC domain), which are found in helicases of multiple families (Linder, 2006) (Fig. 3-1). Also present is an HRDC (helicase and RNaseD C-terminal) domain (approximately aa1207-1283). In contrast to the other domains present, the HRDC domain is only in some RecQ helicases. For example, three human RecQ helicases contain an HRDC domain (WRN, BLM and RECQ1), but it is absent in human RECQ4 and RECQ5 (Bernstein *et al.*, 2010). It is also present in some RNase D homologues (Morozov *et al.*, 1997) and is probably involved in DNA binding (Bachrati & Hickson, 2003). Some RecQ helicases also contain an RQC (RecQ C-terminal) domain found only in this family. The RQC domain has been suggested to be involved in protein-protein interactions (Bernstein *et al.*, 2010), as well as to bind and unwind dsDNA at branch points (Kitano *et al.*, 2010).

However, it is absent from or highly diverged in some members (Bachrati & Hickson, 2003).



**Figure 3-1 Representation of predicted protein domains in TbRECQ2**  
Protein domains predicted from the predicted amino acid sequence of TbRECQ2 using InterProScan 5 and comparison with those of *E. coli* and *S. cerevisiae* RecQ proteins. Sequence length is indicated on the right. Not to scale.

A multiple sequence alignment of the RECQ2 sequence and other eukaryotic RecQ proteins was performed (ClustalX 2.0). Figure 3-2 shows this alignment, with only the central conserved protein core shown for reasons of space. The RecQ protein family, as in all SF2 helicases, is characterised by seven highly conserved protein motifs (I, Ia and II-VI) (Bachrati & Hickson, 2003; Guo *et al.*, 2007). As shown in Fig.3-2, all of these motifs are present in TbRECQ2, which along with P-BLAST searches and domain identification discussed above, confirms that it is a RecQ helicase family member.

P-BLAST searches described above, which returned mainly BLM homologues, suggested that TbRECQ2 is more closely related to eukaryotic BLM homologues. To investigate this further, a multiple sequence alignment was performed of TbRECQ2 and RECQ proteins from a range of organisms and a phylogenetic tree was constructed from the resulting alignment, shown in Figure 3-3 (using CLC Genomics Workbench 6). As would be expected, TbRECQ2 groups closely with other kinetoplastid RecQ proteins. *T. brucei*, *T. cruzi* and *L. major* each contain two predicted RecQ helicases and TbRECQ2 clusters with one RecQ-like helicase from *L. major* and *T. cruzi*. Beyond this, TbRECQ2 clusters with BLM homologues from several eukaryotes (including SGS1, the *Saccharomyces cerevisiae* BLM homologue). This phylogenetic analysis appears to confirm that TbRecQ2 is closely evolutionarily related to SGS1/BLM. In addition, it indicates that

TbRECQ2 is more similar to SGS1/BLM than the other RecQ-like protein in *T. brucei*, TbRecQ1, which is discussed in Chapter 5.

**Figure 3-2 Multiple sequence alignment of RecQ-like helicases.**

The protein sequences of eukaryotic RecQ-like helicases were aligned generated using ClustalX (2.0) ([clustal.org/clustal2/](http://clustal.org/clustal2/)) and shaded using BoxShade ([ch.embnet.org/software/BOX\\_form.html](http://ch.embnet.org/software/BOX_form.html)) and the central conserved helicase core is shown (region indicated by residue numbers on the left). The seven highly conserved SF2 helicase motifs in the helicase domain are indicated. Black background indicates conserved residues and a grey background indicates similar residues. Tb, *T. brucei*; Hs, *Homo sapiens*; Sc, *S. cerevisiae*; Sp, *S. pombe*. Accession numbers for sequences can be found in Appendix 7.2.

|           |     |  |  |   |   |
|-----------|-----|--|--|---|---|
| ScSGS1    | 652 | SSKTNNGPT----                              | YFWSD----                                | EVLRLHEVFKLPGRPNQLEAVNATLQ----                              | GKDVFLMPTGGGKSLCYQLPAVVKSGKT----                |
| HsRECQ1   | 65  | PAAWNKED----                               | FPWSG----                                | KVKDILQNVFKLEKFRPLQLETINVTMA----                            | GKEVFLVMPTEGGKSLCYQLPALCSDG----                 |
| SpRQH1    | 488 | KVLPMSLDDPMLSPWSK----                      |  | EVLGCLKKHFKHGRKNQLEAINGTIS----                              | GKDVFLMPTGGGKSLCYQLPAVIEGAS----                 |
| HsBLM     | 636 | SRNLKHERFQSLSPHFK----                      |  | EMMKIFHKKFGHFRNTQLEAINAALL----                              | GEDCFILMPTGGGKSLCYQLPACVSPG----                 |
| TbRECQ2   | 432 | IAKASTGEFSGEHYPNSE----                     |  | ELRRTMVDVFGHFRFLQLEINACMA----                               | NRDVFILMPTGGGKSLCYQLPALLPNP----                 |
| HsRECQ5   | 10  | -----                                      | FDPER----                                | RVRSTLKKVFGFDSFKTPLQESATMAVVGKNDVFCMPTEGGKSLCYQLPALLAKG---- | -----   |
| HsWRN     | 514 | NEANEGEEDDDKDFLWPAPNE                      | QVTCCKMYFGHSSFKPVQ                       | KVVIHVSVEE-----   | RRDNVAVMATEGKSLCFQYPPVYVGK----                  |
| TbRECQ1   | 129 | KETPRSNTS----                              | TIPICS----                               | YLLQSHLGVSSGEFERGQEAVALILQ----                              | GESTLAVFPTGNGKTLCFQVPLLVHRLLFQOKLM----          |
| HsRECQ4   | 449 | TVLPYSLGPSGLAETP----                       | AEVFQALEQLGHQAFRPGQ                      | ERAVMRILS----   | GISTLLVLPTGAGKSLCYQLPALLYSRRS----               |
| Motif I   |     |  |  |   |   |
| Motif Ia  |     |  |  |   |   |
| ScSGS1    | 722 | -----                                      | -----                                    | HGTIVISPLISLMQDQVEHLNKN-----                                | IKASMFSSRGTAEQEROTFNLFIN-----                   |
| HsRECQ1   | 133 | -----                                      | -----                                    | FTLVIPLISLMQDQVMDLQKLG-----                                 | ISATMLNASSSKEHVWVHAEMVNKNSELKLIYVTPEKIA-----    |
| SpRQH1    | 563 | -----                                      | -----                                    | RGVTLVISPLLSLMQDQDLHLRKLN-----                              | IPSLPLSGEQPADERQVISFLMAKNVLVKLIYVTPEGLA-----    |
| HsBLM     | 709 | -----                                      | -----                                    | VTVVISPLRSILIVDQVQKLTSLD-----                               | IPATYLTGDKTDEATNIYQLSKKDPYIKLLIYVTPEKIC-----    |
| TbRECQ2   | 505 | -----                                      | -----                                    | AKVTIVISPLVSLIQDQVAYRAYD-----                               | LPAMALTGTQTLDAPEERDLFNEWSSGRIVCTLIYVTPEYFG----- |
| HsRECQ5   | 72  | -----                                      | -----                                    | ITIVISPLIALIQDQVDHLLTLK-----                                | VRVSSLNKSLSAQERKELLADLEREKPOTKILYITPEMAA-----   |
| HsWRN     | 591 | -----                                      | -----                                    | IGLVISPLISLMQDQVLQKMSN-----                                 | IPACFLGSAQSEN-----                              |
| TbRECQ1   | 205 | LWKQQLENVSLPEYTAPCSRFGVVVSPLLSLMADQVERI    | TKDGS                                    | LKAVILSSATSGARDKDLLES                                       | SSPVCNIDLVFLSPEKIV-----                         |
| HsRECQ4   | 524 | -----                                      | -----                                    | PCLTLVVSPLLSLMDDQVSGEPPCL-----                              | KAACIHSGMTRKQESVLQKIRA-----                     |
| Motif II  |     |  |  |   |   |
| ScSGS1    | 785 | ASEQCKRAISRLLYADGKLARIIVVDEAHCVSNWGHDFRPDY | KELKFFKREYPD-----                        | IPMIALTATASEQVRMDI  | IHNDELKE-P-----                                 |
| HsRECQ1   | 196 | KSKMFMSRLLEKAYEARFRTRIAVDEVHCQSOWGHDFRPDY  | KALGILKROFPN-----                        | ASLIGLTATATNHVLTDAOKIL                                      | CLIEK-C-----                                    |
| SpRQH1    | 628 | SNGAITRVLKSILYERKILARIVIDEACHCVSNWGHDFRPDY | KQLGLRDYQG-----                          | IPFMALTATANEIVKKDI  | IINTIRMEN-C-----                                |
| HsBLM     | 772 | ASNRLISTLENLYERKILARFVIDEACHCVSNWGHDFRPDY  | KRMNMLRQKFP-----                         | VPVMAALTATANPRVQKDILTQ                                      | KLILR-P-----                                    |
| TbRECQ2   | 570 | RSDFVQSLKSLENRGLNRFVDEACHCVSNWGHDFRPDY     | KRKAALKHFPQ-----                         | VPIALTATATDVQORDIIQTL                                       | GLHN-A-----                                     |
| HsRECQ5   | 135 | SS-SFOPTLNSLVSRHLLSYLVVDEACHCVSNWGHDFRPDY  | LRGALRSRLGH-----                         | APCVALTATATPOVEDVFAAL                                       | HLKKPV-----                                     |
| HsWRN     | 648 | GNMGLLOQLEADIG-----                        | ITLFAVDEACHISEWGHDFRDSFRKLGSLKTALPM----- | VPIVALTATASSSIREDIVRC                                       | INLRN-P-----                                    |
| TbRECQ1   | 290 | RHRELROVLKGQVHR-----                       | IATLCIDEACHCVSNWAYDFRPSFLYVHC            | VEIDIVLGGDFSPFELCLTATATPSVIDD                               | LOSTFNIR-----                                   |
| HsRECQ4   | 586 | GAGGLPP-----                               | AAQLPPVAFACIDEACHCLSQWSHNFRPCXLRVCKVL    | RERMG-----  | VHCFGLGTATATRRRTASDVAQHFAVAEEP-----             |
| Motif III |     |  |  |   |   |
| ScSGS1    | 864 | VFLKQSFNRNTLYEYVNNK                        | -----                                    | TKNTIFEICDAVKSRFK-----                                      | NQTGIYCHSKSKSCETSQAMQ-RNGIK-----                |
| HsRECQ1   | 275 | FTFTASFNRPNLVYEVROKPSN                     | -----                                    | TEDFIEDIVKLINGRYK-----                                      | GOSGIYCFSQKDSQOVTSTQ-NLGIH-----                 |
| SpRQH1    | 707 | LELKSFNRPNLFYELKPK                         | -----                                    | KDLYTELRYFISNGHL-----                                       | HESGIYCLSRSTSCEQVAALRNNDYGLK-----               |
| HsBLM     | 851 | QVFSMSFNRNHLKYYVLPK                        | -K-----                                  | PKKVAFDCLEWIRKHHP-----                                      | YDSGIYCLSRRECRDTMADTLQ-RDGLA-----               |
| TbRECQ2   | 649 | VSFKGSFNRNHLKYYVQRT                        | -----                                    | SKAGSTVAEIIKKNFPP-----                                      | RSCGIYCLSKKDCCEMAAVLR-KEGIR-----                |
| HsRECQ5   | 214 | AIFKTPCFRANLFYDVQFKELISDP                  | -YGNLKDFCLKALGOEADKGLSGCGIVYCTR          | EACEQLAIELS-CRGVN-----                                      | -----   |
| HsWRN     | 724 | QITCTGDFDRPNLYLEVRKFTG                     | -----                                    | NILQDLQPLVKTSSHWFEFPTIYCP                                   | SKMTQVITGELR-KNLIS-----                         |
| TbRECQ1   | 371 | -TVVVPYQORANLQLEAVSLSSENVTKDSPTTHSDLR      | KEILQTVLELPKPMLVVOTRADADELAKFLSSKL       | GAAQNGKARKKGD   | X-----  |
| HsRECQ4   | 663 | DLHGPAVPVPTNHLSTVSMDR                      | -----                                    | DTDQALLTLQGRKFQNLDSIIYICNRR                                 | EDTERIAALRTCLHAAVVPG-----                       |

Figure 3-2



## Motif V

927 ScSGS1 ---CATVAFGMGIDKPDVRFVYETVP  
 931 HsRECQ1 ---IQVVCATVAFGMGIDKPDVRFVYHHSMS  
 935 SprQH1 ---IQVVATVAFGMGIDKPDVRFVYHHSFP  
 939 HsBLM ---YKIVATVAFGMGIDKPDVRFVYHHSFP  
 943 TbRECQ2 ---QVVCATVAFGMGIDKPDVRFVYHHSFP  
 947 HsRECQ5 ---QVVCATVAFGMGIDKPDVRFVYHHSFP  
 951 HsWRN ---QVVCATVAFGMGIDKPDVRFVYHHSFP  
 955 TbRECQ1 ---QVVCATVAFGMGIDKPDVRFVYHHSFP  
 959 HsRECQ4 ---QVVCATVAFGMGIDKPDVRFVYHHSFP

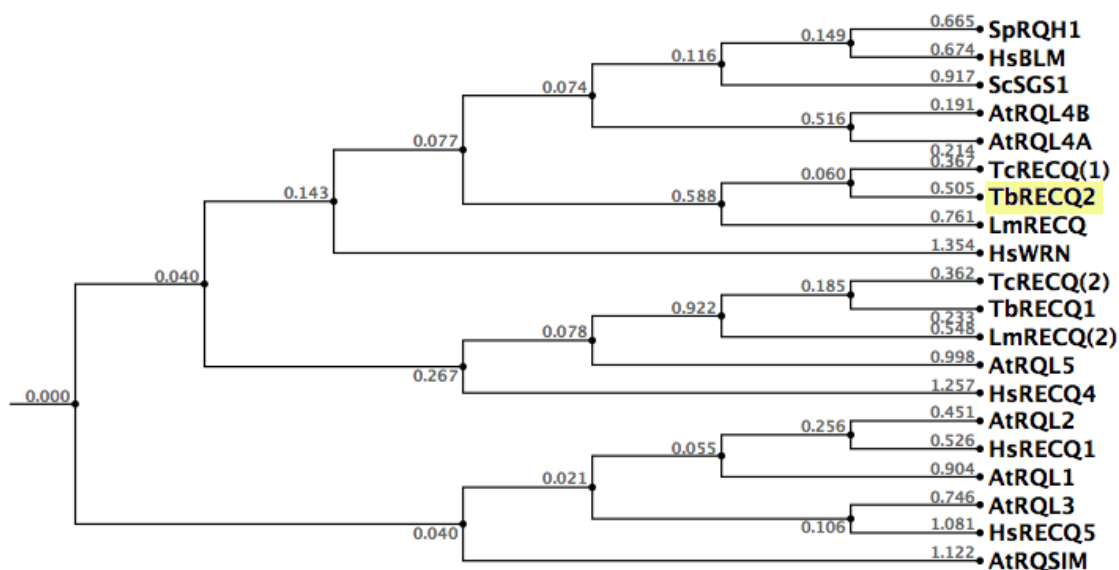
## Motif VI

979 ScSGS1 ---KDKNLDRENKEKHLNKLQVMAVCDNVTDCCRKLVLST  
 983 HsRECQ1 ---KDKNLDRENKEKHLNKLQVMAVCDNVTDCCRKLVLST  
 987 SprQH1 ---KDKNLDRENKEKHLNKLQVMAVCDNVTDCCRKLVLST  
 991 HsBLM ---KDKNLDRENKEKHLNKLQVMAVCDNVTDCCRKLVLST  
 995 TbRECQ2 ---KDKNLDRENKEKHLNKLQVMAVCDNVTDCCRKLVLST  
 999 HsRECQ5 ---KDKNLDRENKEKHLNKLQVMAVCDNVTDCCRKLVLST  
 1003 HsWRN ---KDKNLDRENKEKHLNKLQVMAVCDNVTDCCRKLVLST  
 1007 TbRECQ1 ---KDKNLDRENKEKHLNKLQVMAVCDNVTDCCRKLVLST  
 1011 HsRECQ4 ---KDKNLDRENKEKHLNKLQVMAVCDNVTDCCRKLVLST

1056 ScSGS1 ---EPAKKIVKLVESIQ-----NERVYIIYQDVFVKGS  
 1060 HsRECQ1 ---EYCRDLIKILKQAE-----LNEKLTPLKIDSWMK  
 1064 SprQH1 ---EYCRDLIKILKQAE-----LNEKLTPLKIDSWMK  
 1068 HsBLM ---EYCRDLIKILKQAE-----LNEKLTPLKIDSWMK  
 1072 TbRECQ2 ---EYCRDLIKILKQAE-----LNEKLTPLKIDSWMK  
 1076 HsRECQ5 ---EYCRDLIKILKQAE-----LNEKLTPLKIDSWMK  
 1080 HsWRN ---EYCRDLIKILKQAE-----LNEKLTPLKIDSWMK  
 1084 TbRECQ1 ---EYCRDLIKILKQAE-----LNEKLTPLKIDSWMK  
 1088 HsRECQ4 ---EYCRDLIKILKQAE-----LNEKLTPLKIDSWMK

1117 ScSGS1 ---KIVQANHDTLEEHGICKSMQKSEIERIFFHLITIRVLQEYSIMNNSGFASSVVKVCPNAKKLLTGKMEIK-----MQ  
 1121 HsRECQ1 ---KIVQANHDTLEEHGICKSMQKSEIERIFFHLITIRVLQEYSIMNNSGFASSVVKVCPNAKKLLTGKMEIK-----MQ  
 1125 SprQH1 ---KIVQANHDTLEEHGICKSMQKSEIERIFFHLITIRVLQEYSIMNNSGFASSVVKVCPNAKKLLTGKMEIK-----MQ  
 1129 HsBLM ---KIVQANHDTLEEHGICKSMQKSEIERIFFHLITIRVLQEYSIMNNSGFASSVVKVCPNAKKLLTGKMEIK-----MQ  
 1133 TbRECQ2 ---KIVQANHDTLEEHGICKSMQKSEIERIFFHLITIRVLQEYSIMNNSGFASSVVKVCPNAKKLLTGKMEIK-----MQ  
 1137 HsRECQ5 ---KIVQANHDTLEEHGICKSMQKSEIERIFFHLITIRVLQEYSIMNNSGFASSVVKVCPNAKKLLTGKMEIK-----MQ  
 1141 HsWRN ---KIVQANHDTLEEHGICKSMQKSEIERIFFHLITIRVLQEYSIMNNSGFASSVVKVCPNAKKLLTGKMEIK-----MQ  
 1145 TbRECQ1 ---KIVQANHDTLEEHGICKSMQKSEIERIFFHLITIRVLQEYSIMNNSGFASSVVKVCPNAKKLLTGKMEIK-----MQ  
 1149 HsRECQ4 ---KIVQANHDTLEEHGICKSMQKSEIERIFFHLITIRVLQEYSIMNNSGFASSVVKVCPNAKKLLTGKMEIK-----MQ

Figure 3-2



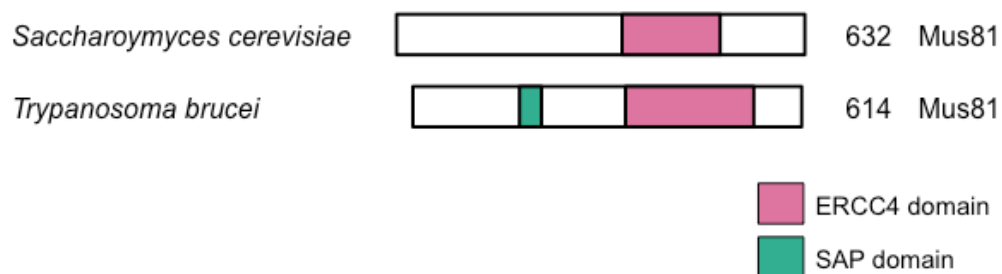
**Figure 3-3 Neighbour-joining tree of eukaryotic RecQ-like helicases**  
Protein sequences from multiple eukaryotic RecQ-like proteins were aligned and a tree generated from this alignment using CLC Genomics Workbench 6, using the neighbour joining method. Hs, *H. sapiens*; Sc, *S. cerevisiae*; Tb, *T. brucei*; Sp, *S. pombe*; At, *Arabidopsis thaliana*; Lm, *Leishmania major*; Tc, *Trypanosoma cruzi*. Accession numbers for sequences can be found in Appendix 7.2.

### 3.2.2 Sequence analysis of *T. brucei* MUS81

Initial identification of TbMUS81 was through P-BLAST analysis (NCBI nr November 2014) using the *S. cerevisiae* MUS81 sequence against the *T. brucei* Lister 427 proteome (using tritrypdb.org; Version 8.1), which returned only the TbMUS81 sequence (TriTrypDB accession number Tb427.08.6740), though with an expectancy value of only 0.094. Further analysis was therefore required to confirm the identity of the putative TbMUS81.

As for TbRECQ2, protein P-BLAST analysis using the predicted sequence of TbMUS81 as the query was performed to investigate the similarity of TbMUS81 to other eukaryotic MUS81 proteins. The top hits were DNA repair and hypothetical proteins from *Trypanosoma* and other kinetoplastid species, including *Trypanosoma vivax*, *T. cruzi*, *Leishmania mexicana*, *Leishmania infantum*, *L. major* and *Phytomonas spp.* Non-*Trypanosoma spp.* Hits retrieved were dominated by fungal, parasitic and plant hypothetical proteins, some of which were annotated as MUS81-like. Upon restricting the search to *S. cerevisiae* or *H. sapiens*, MUS81 proteins from both species were returned, with expectancy values of 0.52 and  $10^{-4}$ , respectively. Protein domain analysis using InterProScan

sequence search (<http://www.ebi.ac.uk/interpro/search/sequence-search> (Hunter *et al.*, 2012)) was used to analyse the domains present in MUS81, which are shown in Figure 3-4. This revealed a predicted ERCC4 domain (approximately aa329-523), with an expectancy value of 0.02. ERCC4 domains are involved in nucleic acid binding and are found in a number of DNA repair proteins, including RAD1 and XPF, as well as in MUS81 and another crossover junction endonuclease, EME1 (Ciccia *et al.*, 2008). Additionally, a SAP (SAF-A/B, Acinus and PIAS) domain was predicted near the N terminus of TbMUS81 (aa164-198). It has been proposed that SAP domains are involved in DNA binding (Aravind & Koonin, 2000), though they are not found in other characterised eukaryotic MUS81 proteins such as in *M. musculus* and *H. sapiens*.



**Figure 3-4 Representation of protein domains predicted in TbMUS81**  
Protein domains predicted from the predicted amino acid sequence of TbMUS81 using InterProScan 5. Sequence length is indicated on the right. Not to scale.

The low similarity of TbMUS81 to other characterised eukaryotic MUS81 proteins, the lower expectancy value for the MUS81-characteristic ERCC4 domain and the unusual SAP domain was the reason for then using PSI-BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) search to attempt to confirm the identity of Tb427.08.6740 as encoding TbMUS81. TbMUS81 was used as the query sequence and convergence was reached on the 16<sup>th</sup> iteration. The top 10 hits were primarily fungal MUS81 and putative MUS81 homologues, ERCC4 domain-containing proteins and endonuclease proteins. Additionally, P- BLAST searches using Mus81 sequences from eukaryotes including *H. sapiens* and *Arabidopsis thaliana* as the query against the *T. brucei* genome (data not shown) (<http://tritrypdb.org>) retrieved *T. brucei* MUS81 as the only significant alignment in the *T. brucei* genome.

A multiple sequence alignment was performed (ClustalX 2.0) of TbMUS81 and eukaryotic MUS81 homologues to further investigate the similarity of TbMUS81 to these proteins (Fig. 3-5). Overall, the average sequence identity of MUS81 to

the MUS81 sequences in Figure 3-5 is only 11.5%, though it is higher (17.2%) within the ERCC4 domain. MUS81 proteins contain an ERKx3D motif in the ERCC4 domain (Chang *et al.*, 2008), which is also present in TbMUS81 (Fig. 3-5).

**Figure 3-5 Multiple sequence alignment of MUS81 proteins**  
The protein sequences of multiple eukaryotic MUS81 proteins were generated using ClustalX (2.0) ([clustal.org/clustal2/](http://clustal.org/clustal2/)) and shaded using BoxShade ([ch.embnet.org/software/BOX\\_form.html](http://ch.embnet.org/software/BOX_form.html)). Black background indicates conserved residues and a grey background indicates similar residues. Tb, *T. brucei*; Hs, *Homo sapiens*; Sp, *S. pombe*; YEAST, *S. cerevisiae*. Accession numbers for sequences can be found in Appendix 7.2.

|         |     |                    |  |
|---------|-----|--------------------|--|
| HsmUS81 | 1   | MAAPVRLGRKRPPEA    | CPNPLFVRMLTEWRDEATRSRRRTFRVFOKALRSLRRYPPLPLRSKKEAKIQHFGDGLCRMEDERLQRHRTSGGDHAPDS               |
| SpmUS81 | 1   | -----              | -----MDCCNPLFLQMIQEWMEESTRRFPKSYQTRWKAYDSMKSCPIITHRPSQALAKIGIGPTCAKDEKKNAYCLENNIPISTH          |
| ScmUS81 | 1   | -----              | -----MESSNLKDLVYIEMLOELVDGLTFPKQEQLKIAVEKAKRNQNAEGSFYPTDLKVKIGICNLIIRKLTDKIRNYCKIHILSPVEA      |
| TbmUS81 | 1   | -----              | -----MNSHNRIASIIYERRWPQR-----PFEFCAALRRYPPLPISSGAAGTQFHGSCFEAVFATEAFCAMLAPRHGREVLGD            |
| HsmUS81 | 96  | PSGENSPAQGR        | -----LAEQVDSMP--VPAQPKAGGSGSYMPARHSGARVILLVLYREHLNEN--GHFLLTKEELLQRCQAQKSPRVAPG                |
| SpmUS81 | 83  | NEONDSHVNANK       | -----SSSETSEKPRSVKPTTRKRKYVPSYRSGAYSILCALAMNKKHEF--ATKPOLVTMAQPVCDSSFGSATDR                    |
| ScmUS81 | 86  | PSLTQTSSTRPPKRTT   | TALRSIVNSCENDKNEAPEEKGTKRKRKYIPKRRSGGYAILSLLENAIERGVSKQELIEVAGKYSDHCMTPNFST                    |
| TbmUS81 | 74  | DNSTENDEMLARDGEVSX | -----ALRRSVSPFSPFCFKMRRTGCGPPLLSQPAVWSGNSDWHFGOLHNSRGEAGTSLMHSSTTGCQSGEPL                      |
| HsmUS81 | 178 | SAR--PMPALRS       | LHRNIAVLRTHOPARYSLTPEGLELAOKLAESEG-----LSLLNVGIGPKE-----PPGEETAVPGAAS-----                     |
| SpmUS81 | 166 | NMRYTANSAMKTL      | ITKNIAVYQTHPSKYCLTDGEEVCIRLAKVDDSFQRK--HTVSNFSVSKSDHSDSLCQDPENFVTSINKAGSSSDHGGELH              |
| ScmUS81 | 181 | KEFYGAMSSIAALK     | KHSLVIEEGRPKRYSLTEEGVELTKSLKTADGISFPKENEENPEYSVTNESESSEETANLTDLRGEYGKEEPCDINNTSFM              |
| TbmUS81 | 163 | PEEKLSTRETRDRCLS   | KGLTGSG-TKVHLIARLRKHGASKLAETSPKPRYPVPSQAEVLSHLREDSEKSRSMSSAGHTSTVSDMIRHGQRLRTL                 |
| HsmUS81 | 247 | -----              | -----AELASEAGVQ-----QOPLRLRPGEYR-----VLTQVDIGETRGCGHRPELLR                                     |
| SpmUS81 | 259 | VTYCPVDHNEVSDGVE   | TDIDVDQVDSLTGIGHDHIINN--EQLIDLTEQEKQKPNESSLNLKIEVLFSECTVFLIIDTREIRSPLDRLIID                    |
| ScmUS81 | 276 | LDITFQDLSTPQR      | QNNVFNKDRHNSQTNISSHKLEEVDDQTPVDSALKAKSTIKRRRYNGVSYELWCGDFEVPIIDHREIKSQSDREFFSR                 |
| TbmUS81 | 257 | FRSPNTDIAPTSTAE    | GGSEPGRLLGTPKEFVASQQRERKXVHSTPSMPTNTDIDADENETPKKEGCSNYSWHLLVDNRERIKG-AHENLIE                   |
| HsmUS81 | 290 | ELQR-LHVTHTRK      | HHVGDFFVVAQETNPRDPANPG-----ELVLDHIVERKRLDDLCSSIIDGRFREOKFRLKR                                  |
| SpmUS81 | 351 | KLTFNDFGNCVRS      | DELGDALVARDMESGQ-----EVLDFVVERKRDLDLVAISKDGRFHEOKARLKK   |
| ScmUS81 | 371 | AFER-KGMSERQAL     | GLDILINVAKNKNTGL-----QCVLNTIVERKRLDDLAISIRDNRFMEOKNRLEK  |
| TbmUS81 | 351 | VCSR-ACAPSVSCTLP   | CGDFMIGINSPNVGCSSANSAXVPDYQCRNAAGCTVLRGAMYGTQKISFLVVERTVKDKCASITSSRYVEQORQLSS                  |
| HsmUS81 | 359 | CGLERRVIAVEEHGS    | VHNLISLPESLTLLQAVTNTQVIGFFVKRTADIKESAAYLALLTRGLQRLYQG-HTTRSRP-----WGTPGNPESGAMTS-              |
| SpmUS81 | 416 | SGIRSVTHIEES--     | SYDESFTES-IRTAVSNTQVDOLPHVRHRTSRLEHSVSLAEEMKQINLFEKRETLAVIPDLSIEAKTYESLREQLKI-                 |
| ScmUS81 | 435 | SGCEHKYTLIEENSG    | NIIGNMNEA-LIKTALWVILVYKFSMIRTCNSDETVEKIHATHTVSHHSQKDLIVIFPS---DLKSKDDYKKVLLQFR                 |
| TbmUS81 | 445 | SPFRSVVWVEGEMES    | ITVEERRR-VLSACASLASLPRFRVUVWTRHSETATFRSLGKSVGSILCK-----TKMEQCPETLVTN-                          |
| HsmUS81 | 447 | -----              | -----PNPLCSLLTFSDFNAGAIKNKAQSVREVFARQLMQVRGVSGEKAALVDRYSTPASLAAVADACATPKQEQTLLS-TIKCGRRLQRLNLP |
| SpmUS81 | 507 | -----              | -----DPSTPYHISYHAFSSVLSKSTLTGVGDIFIRMLMTIKGISASKAIEIOKKYPTFMHFFEAVERSSSQERNMLIN-KTCQGYGFOTIGPA |
| ScmUS81 | 526 | REFERKGGIECHNLE    | CFQELMGDGLTKVGELTTHVLMVKGISLEKAVIOEIFPTLNKIMAYKCSSEEEAKLLMNFNLGADAPGAKKITKS                    |
| TbmUS81 | 524 | -----              | -----CTECLQYINKIRKDIQARTTFPRMLMCIRGCSAPLAIQLASKYGSLLGFWRLRYQG--VDACDADPDIHRLTKPOKEVYLC         |
| HsmUS81 | 536 | LSRTLSQLYCSYGPLT   |  |
| SpmUS81 | 596 | LSAKVASVTFPES      | -----  |
| ScmUS81 | 621 | LSEKTYDALGKIL      | -----  |
| TbmUS81 | 604 | LTAKFLTRSYC        | -----  |

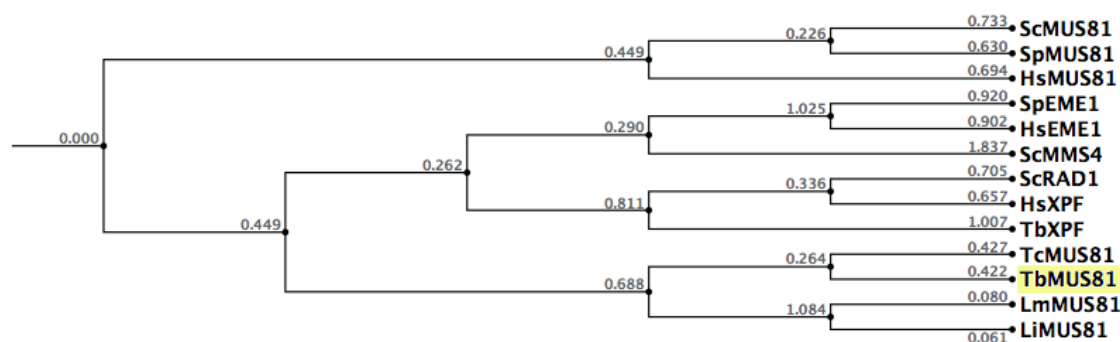
## Conserved ERKX3D Motif

## ERC4 domain

Figure 3-5



ERCC4 domains are present in a number of DNA repair nucleases in addition to MUS81, including XPF, RAD1 and EME1/MMS4. To confirm that TbMUS81 is a MUS81 homologue rather than another ERCC4-containing protein, its similarity to other ERCC4 proteins was analysed. An alignment of eukaryotic MUS81, EME1/MMS4 and XPF proteins was generated and used to construct a tree using CLC Genomics Workbench 6 (Fig. 3-6). Unsurprisingly, TbMUS81 groups closely with other kinetoplastid (predicted) MUS81 proteins. However, all of these MUS81 proteins group more closely with EME1/MMS4 and XPF proteins than to other eukaryotic MUS81 proteins. It appears therefore that *T. brucei* MUS81, along with other kinetoplastid MUS81 proteins, is much diverged from yeast and human MUS81 proteins.



**Figure 3-6** Neighbor-joining tree of eukaryotic ERCC4-containing proteins. Protein sequences from multiple eukaryotic EME1/MMS4, RAD1, MUS81, and XPF proteins were aligned and a tree generated from this alignment using CLC Genomics Workbench 6, using the neighbour joining method. Hs, *H. sapiens*; Sc, *S. cerevisiae*; Tb, *T. brucei*; Sp, *S. pombe*; Li, *Leishmania infantum*; Lm, *Leishmania major*; Tc, *Trypanosoma cruzi*. Accession numbers for sequences can be found in Appendix 7.2.

Taken together, these analyses indicate that TbMUS81 appears to be a MUS81 homologue, though as with many kinetoplastid proteins, very much diverged from other eukaryotes.

### 3.2.3 TbPIF6 is a member of the PIF1-like helicase family

Searches conducted using P-BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>; NCBI nr November 2014) with *T. brucei* Lister 427 PIF6 (TriTrypDB accession number Tb427.10.910) (Liu *et al.*, 2009a) as the query retrieved known Pif1-family helicases as good matches, including PIF1 helicases of *H. sapiens*, *M. musculus* and *Xenopus laevis* at expectancy values of approximately  $10^{-70}$ . The top hits however, excluding kinetoplastids, were predicted PIF1-like proteins from *Anopheles spp.* and a number of vertebrate species.

Additionally, reciprocal P-BLAST searches on tritrypdb.org (Version 8.1) using these sequences as the query sequences retrieved all TbPIF1 helicases including PIF6. Protein domain analysis using InterProScan sequence search (<http://www.ebi.ac.uk/interpro/search/sequence-search/>) (Hunter *et al.*, 2012) revealed TbPIF6 to contain an AAA+ ATPase domain (approximately aa236-388) (Fig. 3-7). Other PIF1 proteins contain at least one of these domains or another domain of the P-loop NTPase (nucleotide triphosphatase) family involved in nucleotide binding. Additionally, the TbPIF6 sequence was identified by InterProScan as a PIF1-like helicase.



**Figure 3-7 Representation of protein domains predicted in TbPIF6**  
Protein domains predicted from the predicted amino acid sequence of TbPIF6 using InterProScan 5. Approximate position (from the N-terminus) is shown underneath the diagram. Not to scale.

A multiple protein sequence alignment was constructed of TbPIF6 with Pif1-like helicases from a number of eukaryotic species including other kinetoplastids (Fig. 3-8). TbPIF6 is well aligned with other Pif1-like family helicases, though the N and C termini of the sequences are less well aligned and are not shown in the alignment in Figure 3-8. TbPIF6 contains all of the seven SF1s (I, Ia, II-VI) motifs (Fig. 3-8), as well as the putative Pif1-like family signature, located between motifs II and III, identified by Bochman *et al.* (2010) in their analysis of Pif1-family helicases.

**Figure 3-8 Multiple sequence alignment of Pif1 family helicases**  
The protein sequences of multiple eukaryotic Pif1 family helicases were aligned and a tree generated using ClustalX (2.0) ([clustal.org/clustal2/](http://clustal.org/clustal2/)) and shaded using BoxShade ([ch.embnet.org/software/BOX\\_form.html](http://ch.embnet.org/software/BOX_form.html)). Black background indicates conserved residues and a grey background indicates similar residues. The conserved helicase core is shown and the seven SF1 helicase motifs (grey arrows) and the putative Pif1 family signature (orange arrow) are shown. Sc, *S. cerevisiae*; Sp, *S. pombe*; Hs, *Homo sapiens*; Tb, *T. brucei*. Locations for helicase motifs I-VI and putative Pif1 helicase family signature are indicated. Accession numbers for sequences can be found in Appendix 7.2.

|        |     |                    |       |       |                    |        |                |          |       |       |      |        |       |       |       |       |       |     |       |       |       |       |       |       |   |   |   |   |   |       |   |   |   |   |   |   |   |       |       |       |       |   |   |       |       |       |       |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|--------|-----|--------------------|-------|-------|--------------------|--------|----------------|----------|-------|-------|------|--------|-------|-------|-------|-------|-------|-----|-------|-------|-------|-------|-------|-------|---|---|---|---|---|-------|---|---|---|---|---|---|---|-------|-------|-------|-------|---|---|-------|-------|-------|-------|---|---|-------|---|---|---|---|---|---|---|---|---|---|---|---|---|-------|---|---|---|---|-------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ScRRM3 | 1   | -----MFRSHASGNKKQW | KNS   | SSNGN | TPAASASGSHAYRQQTLS | SFFMGS | GKKSAAAAKSDSTI | IDL      | ES    | ----- |      |        |       |       |       |       |       |     |       |       |       |       |       |       |   |   |   |   |   |       |   |   |   |   |   |   |   |       |       |       |       |   |   |       |       |       |       |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| ScPIF1 | 1   | -----MPKWIRSTLNHI  | IPRRP | FI    | CS                 | SF     | FLLLK          | NVSHAKLS | FSSSR | GER   | SNNF | IOAQLK | HP    | SIL   | SKED  | LD    | LL    | SDS | ----- |       |       |       |       |       |   |   |   |   |   |       |   |   |   |   |   |   |   |       |       |       |       |   |   |       |       |       |       |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| HsPIF1 | 1   | -----MLSGIEAAAGEY  | ED    | SEL   | RCV                | AVE    | EL             | SPGQ     | RRR   | QAL   | RT   | AE     | LS    | GR    | NER   | REL   | ML    | R   | LQAP  | ----- |       |       |       |       |   |   |   |   |   |       |   |   |   |   |   |   |   |       |       |       |       |   |   |       |       |       |       |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| SpPFH1 | 1   | MFSCQSLYKFS        | HRK   | RI    | PV                 | QRA    | QK             | SLL      | HT    | QNE   | SS   | HP     | SL    | NK    | LG    | FS    | SS    | LN  | F     | SS    | QSD   | NL    | PSS   | PIT   | L |   |   |   |   |       |   |   |   |   |   |   |   |       |       |       |       |   |   |       |       |       |       |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| TbPIF6 | 1   | -----MVVLRPKITLT   | SV    | R     | G                  | R      | I              | Q        | V     | T     | X    | E      | D     | G     | E     | R     | V     | G   | P     | W     | G     | T     | E     | C     | F | L | S | R | Q | T     | G | C | P | C | L | V | V | R     | S     | S     | R     | H | R | H     | Q     | G     | T     | F | R | ----- |   |   |   |   |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| ScRRM3 | 65  | -----GDEGNRNIT     | AP    | P     | R                  | L      | R              | N        | S     | S     | S    | L      | F     | S     | Q     | S     | Q     | S   | ----- | F     | G     | D     | D     | D     | P | D | A | E | F | K     | K | L | V | D | V | P | R | ----- |       |       |       |   |   |       |       |       |       |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| ScPIF1 | 73  | -----DDWEEPDCI     | Q     | E     | T                  | E      | K              | Q        | E     | K     | K    | I      | I     | D     | I     | H     | K     | E   | D     | P     | V     | D     | K     | K     | P | M | R | D | K | N     | V | M | N | F | I | N | K | D     | S     | P     | L     | S | W | N     | ----- |       |       |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| HsPIF1 | 61  | -----GPAGRPRC      | F     | P     | L                  | R      | A              | A        | R     | L     | F    | T      | R     | F     | A     | E     | A     | G   | R     | S     | T     | L     | R     | L     | P | A | H | D | T | P     | G | A | V | Q | L | L | S | D     | C     | P     | P     | D | R | ----- |       |       |       |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| SpPFH1 | 86  | PAKGRSAAS          | L     | Q     | L                  | D      | N              | T        | V     | G     | F    | D      | V     | S     | K     | P     | S     | L   | P     | V     | F     | E     | N     | S     | G | L | G | S | K | Y     | S | T | E | A | N | G | V | I     | D     | E     | N     | D | F | D     | D     | D     | L     | L | E | N     | D | I | D | Q | K | P | I | P | W | S | S | P | I | E     | H | T | K | L | T     | K | S | M |   |   |   |   |   |   |   |   |   |   |   |   |
| TbPIF6 | 64  | -----LEGVRQVLS     | L     | Y     | A                  | M      | E              | G        | K     | L     | T    | V      | V     | P     | H     | Q     | R     | L   | C     | S     | V     | F     | I     | E     | T | F | A | D | V | D     | A | L | Q | M | A | A | T | L     | Q     | D     | R     | S | R | ----- | W     |       |       |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| ScRRM3 | 114 | INSYKKS            | RS    | LS    | MT                 | S      | S              | L        | H     | K     | T    | A      | S     | A     | T     | T     | Q     | K   | T     | ----- | Y     | H     | F     | D     | E | T | L | R | E | V     | T | S | V | K | S | N | S | R     | Q     | L     | S     | F | T | S     | T     | I     | N     | I | E | D     | S | S | M | K |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| ScPIF1 | 126 | DMFKPS             | I     | I     | O                  | P      | P              | Q        | L     | I     | S    | E      | N     | S     | F     | D     | S     | S   | Q     | K     | S     | R     | ----- | S     | T | G | F | K | N | F     | L | P | A | L | K | K | E | S     | F     | D     | E     | L | Q | N     | N     | S     | I     | S | Q | E     | R | S | L | E | M | I |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| HsPIF1 | 113 | RRFLRT             | L     | R     | L                  | K      | L              | A        | A     | P     | G    | P      | A     | S     | A     | R     | A     | Q   | L     | L     | G     | ----- | P     | R     | P | R | D | F | V | T     | I | S | P | V | Q | P | E | E     | R     | R     | L     | R | A | A     | T     | R     | V     | P | D | T     | L | V |   |   |   |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| SpPFH1 | 171 | ESSEKRS            | K     | N     | H                  | L      | S              | K        | I     | Y     | E    | D      | H     | T     | S     | E     | K     | G   | A     | S     | S     | V     | I     | S     | S | N | I | T | R | Q     | I | K | R | S | T | L | P | W     | A     | D     | P     | Y | R | Y     | G     | D     | P     | D | E | K     | R | T | S | T | S | A | D | I | S | Q | H | T | V | S     | N | D | S | S | N     | K | L | S | N | G | R | S |   |   |   |   |   |   |   |   |
| TbPIF6 | 118 | KDIEKN             | V     | A     | C                  | R      | V              | Q        | R     | I     | T    | R      | N     | V     | D     | T     | K     | R   | I     | T     | N     | M     | D     | ----- | C | G | T | D | H | L     | D | Y | E | Q | F | E | W | G     | G     | E     | D     | D | S | C     | A     | V     | G     | G | V | S     | C | N | P | L | E |   |   |   |   |   |   |   |   |       |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| ScRRM3 | 179 | LSTDSE             | R     | P     | A                  | K      | R              | S        | K     | P     | S    | M      | E     | F     | Q     | L     | K     | L   | T     | V     | P     | K     | K     | I     | K | L | P | L | R | K     | T | V | S | N | M | D | S | M     | N     | H     | R     | S | A | S     | S     | P     | ----- | V | V | L     | T | M | E | Q | E | R | V | V | N | L | I | V | K | K     | R | T | N | V | F     | Y | T | G | S | A | G | T | G | K | S | V |   |   |   |   |
| ScPIF1 | 189 | NE                 | N     | E     | K                  | K      | M              | O        | F     | G     | E    | K      | I     | A     | V     | L     | T     | Q   | R     | P     | S     | F     | T     | E     | L | Q | N | D | Q | ----- | D | S | N | L | N | F | H | N     | G     | V     | K     | V | K | I     | P     | ----- | I     | C | L | S     | K | E | O | E | S | I | I | K | L | A | E | N | C | ----- | H | N | I | F | Y     | T | G | S | A | G | T | G | K | S | I |   |   |   |   |   |
| HsPIF1 | 173 | KRP                | V     | E     | P                  | O      | A              | G        | A     | E     | P    | S      | T     | E     | A     | P     | R     | W   | P     | ----- | ----- | L     | E     | V     | K | R | L | S | L | P     | S | T | K | P | Q | L | S | E     | O     | A     | A     | V | L | R     | A     | V     | L     | K | C | ----- | Q | S | L | F | F | T | G | S | A | G | T | G | K | S     | Y |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| SpPFH1 | 256 | SLD                | S     | L     | A                  | K      | K              | R        | M     | S     | K    | S      | T     | P     | Q     | I     | S     | K   | K     | F     | S     | V     | P     | L     | N | S | A | S | K | S     | P | I | G | S | S | L | F | K     | T     | S     | D     | S | R | K     | K     | S     | V     | S | I | F     | L | S | D | E | Q | K | R | I | L | D | M | V | E | Q     | Q | H | S | I | F     | F | T | G | S | A | G | T | G | K | S | V |   |   |   |   |
| TbPIF6 | 180 | G                  | A     | R     | G                  | E      | R              | N        | G     | D     | G    | V      | E     | S     | I     | L     | Q     | E   | R     | P     | T     | G     | L     | P     | S | A | H | S | G | G     | K | L | P | K | H | M | G | G     | E     | L     | N     | E | Q | L     | E     | G     | S     | T | T | P     | G | R | M | Q | W | T | D | Q | I | V | A | T | R | L     | V | C | T | G | ----- | S | N | V | F | I | T | G | S | A | G | T | G | K | T | E |
| ScRRM3 | 263 | TLQ                | T     | I     | R                  | Q      | L              | S        | S     | L     | T    | G      | K     | E     | S     | ----- | I     | A   | I     | T     | A     | S     | T     | G     | L | A | A | V | T | I     | G | G | S | T | L | H | K | W     | S     | G     | I     | G | I | G     | N     | K     | T     | I | D | Q     | L | V | K | I | Q | S | Q | K | D | L | L | A | A | W     | R | Y | T | K | V     | L | I | I | D | E | I | S | M | V | D | G | N |   |   |   |
| ScPIF1 | 267 | LR                 | E     | M     | I                  | K      | V              | L        | K     | G     | I    | V      | G     | R     | E     | N     | ----- | V   | A     | V     | T     | A     | S     | T     | G | L | A | A | C | N     | I | G | I | T | L | H | S | F     | A     | G     | I     | G | L | G     | K     | G     | D     | A | D | K     | L | V | K | V | R | R | S | R | K | H | L | R | M | E     | N | I | G | A | L     | V | V | D | E | I | S | M | L | D | A | E |   |   |   |   |
| HsPIF1 | 237 | LL                 | K     | R     | I                  | L      | G              | S        | L     | P     | P    | T      | G     | ----- | ----- | T     | V     | A   | T     | A     | S     | T     | G     | V     | A | A | C | H | I | G     | G | T | L | H | A | F | A | G     | I     | G     | S     | Q | A | P     | L     | A     | Q     | R | P | ----- | G | V | R | Q | G | W | L | N | C | Q | R | L | V | I     | D | E | I | S | M     | V | E | A | D |   |   |   |   |   |   |   |   |   |   |   |
| SpPFH1 | 341 | LR                 | K     | I     | I                  | E      | V              | L        | K     | S     | K    | R      | Q     | S     | D     | R     | V     | A   | T     | A     | S     | T     | G     | L     | A | A | C | N | I | G     | G | T | L | H | S | F | A | G     | V     | L     | A     | R | E | S     | V     | D     | L     | V | S | I     | K | K | N | K | C | V | N | R | M | L | R | T | R | V     | L | I | I | D | E     | I | S | M | V | D | A | E |   |   |   |   |   |   |   |   |
| TbPIF6 | 264 | WL                 | L     | H     | L                  | V      | R              | N        | V     | L     | P    | R      | D     | R     | ----- | ----- | T     | V   | V     | T     | A     | S     | T     | G     | M | S | A | R | L | L     | G | G | C | T | I | H | S | F     | A     | G     | I     | G | E | G     | G     | F     | N     | R | V | N     | R | V | N | R | V | K | S | K | P | E | V | V | R | A     | W | R | Q | C | T     | L | I | I | D | E | I | G | N | I | S | P | D |   |   |   |
| ScRRM3 | 346 | LD                 | K     | L     | E                  | O      | I              | A        | R     | R     | I    | R      | ----- | ----- | K     | N     | D     | P   | F     | G     | G     | I     | Q     | L     | V | L | T | G | D | F     | F | Q | L | P | P | V | A | K     | ----- | ----- | K     | D | E | H     | N     | V     | V     | K | F | C     | F | E | S | E | M | M | K | R | C | I | K | K | T | I     | L | I | L | T | K     | V |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| ScPIF1 | 350 | LD                 | K     | L     | D                  | F      | I              | A        | R     | K     | I    | R      | ----- | ----- | K     | N     | H     | Q   | P     | F     | G     | G     | I     | O     | L | I | F | C | G | D     | F | F | Q | L | P | P | V | S     | K     | ----- | ----- | D | P | N     | R     | P     | ----- | T | K | F     | A | F | E | S | K | A | M | K | E | G | V | K | M | T     | I | M | L | Q | K     | V |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| HsPIF1 | 315 | LD                 | K     | L     | E                  | A      | V              | A        | R     | V     | R    | -----  | ----- | Q     | Q     | N     | K     | P   | F     | G     | G     | I     | O     | L     | I | F | C | G | D | F     | F | Q | L | P | P | V | T | K     | ----- | ----- | G     | S | Q | P     | P     | ----- | R     | F | C | F     | O | S | K | S | M | K | R | C | V | P | V | T | L | E     | L | T | K | V |       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| SpPFH1 | 426 | LD                 | K     | L     | E                  | E      | V              | A        | R     | V     | I    | R      | ----- | ----- | K     | D     | S     | K   | P     | F     | G     | G     | I     | O     | L | V | L | T | G | D     | F | F | Q | L | P | P | V | P     | E     | ----- | ----- | N | G | K     | E     | ----- | S     | K | F | C     | F | E | S | Q | T | M | K | S | A | L | D | F | T | I     | G | L | T | H | V     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| TbPIF6 | 345 | TF                 | S     | M     | I                  | D      | E              | I        | A     | R     | S    | L      | R     | G     | A     | P     | E     | K   | P     | F     | G     | G     | I     | Q     | L | V | L | T | G | D     | F | L | Q | L | P | P | V | D     | S     | P     | K     | A | R | N     | E     | W     | T     | N | G | N     | D | T | D | T | S | N | P | I | P | G | K | L | K | W     | C | F | E | T | A     | T | W | E | S | L | K | L | A | L | V | G | F | R | K | S |

Motif I

Motif II

Motif Ia

Motif IV

Putative Pif1 family signature Motif III

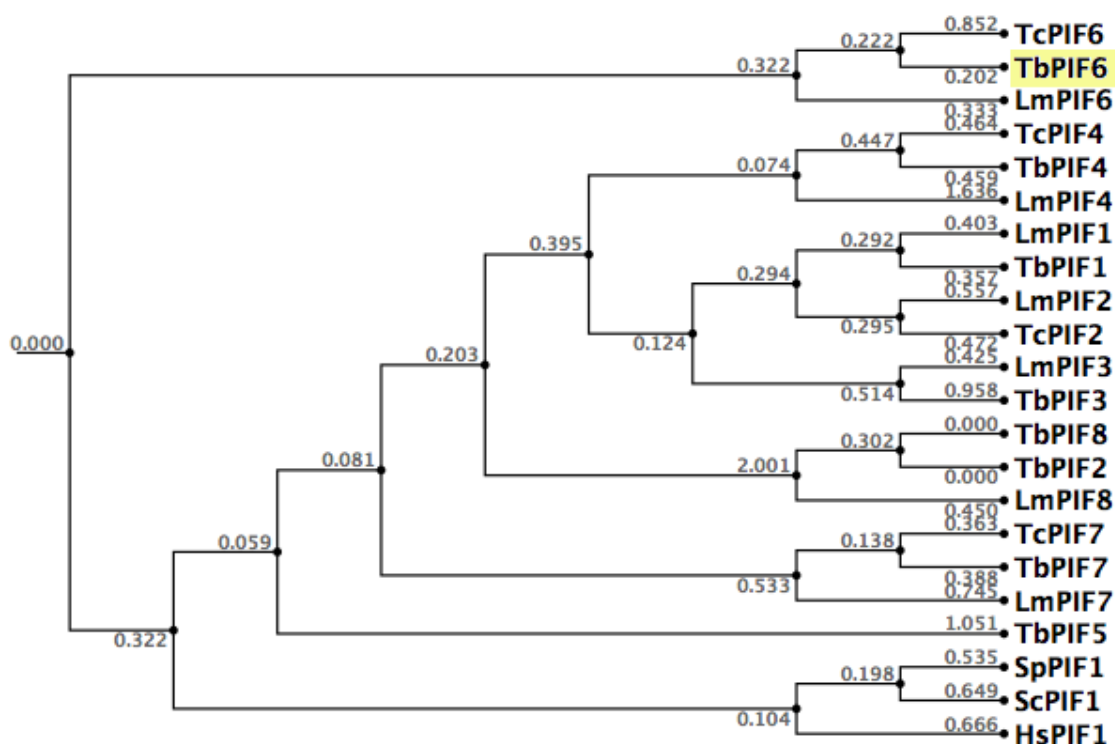
Figure 3-8



|        |     |   |
|--------|-----|---|
| SCRRM3 | 413 | FRQQ-DNKLIDILNAIRYGELTVDAIAKTIRNEN-RDIDYADGGIAPTELYATRRVELSNVKKLQSLP-GDLYEFKAVDNAPER--    |
| SCPIF1 | 416 | FRORGDVKEIDMLNRMRLGNIIDDETEREFKKLS-RPLPDDE-IIIPAEIYSRMEVERANNSRLSKLP-QQVHIFNAIDGGALEDE    |
| HSPIF1 | 380 | WROA-DQTFISLLQAVRLGRCSDEVTRQLQATA-SHKVGRDGIIVATRLCTHODDVALTNERRLOELP-CKVHRFEAMD--NP       |
| SPPFH1 | 491 | FRQK-DEEFVKMLNELRLGKLSDESVRKFVLEN-RTTEYEDGLLPTELPFTRYEVERSNDMRQQIN-QNPVTFTAIDSGTVRDK      |
| TbPIF6 | 430 | VRQMDPDDEALCLEDIRFGRYTRRVERILNECSTRQIKERHGIETPLIVARRDEATEYNAERLKMEDVHFRYESEDYAAIPG-       |
| SCRRM3 | 493 | -YQAILDSSLMVEKVVALKEDAQVMMMLKNK-PDVELVNGSLGKVLFFVTESLVVKMKEIYKIV-----                     |
| SCPIF1 | 498 | ELKERLLQNFAPKETELKVGQAQVMMVKN--LDATLVNGSLGKVIFFMDPPETIFYCYEALTNDPSPPEKLETWAENPSKLKAAME    |
| HSPIF1 | 458 | ELASTLDAQCPVSQLLQLKLGAQVMLVKNLSVSRGLVNGARGVVVGFEAE-----                                   |
| SPPFH1 | 573 | EFRDRLLQGCMAPATLVKKVNAQVMLIKN--IDQLVNGSLGKVIIGFIDD-----ETYQME-----                        |
| TbPIF6 | 514 | ---MNLEKEVSLQQLLELRICGAQVVLASLPDAPHLNSGDQGVVVSFAEQTR-----                                 |
| SCRRM3 | 554 | -DDEVMDMRLVSRVIGNPPLLKESKEFRQDLNARP-----LARLERLKILINYAVKISPHKKEKFPYVRMTVGK                |
| SCPIF1 | 581 | REQSDGEESAVASRKSSVKEGFAKSDIGEPVSPLDSSVDFMKNRVKTDDEVVLENIKRKEQLMQTIHQNSAGKRRRLPLVRFKASD    |
| HSPIF1 | 508 | -----   |
| SPPFH1 | 627 | -KKDAEMQGRNAFEYDS-----LDISPFD-----LPDVKKKKYKLIAMRKASSTAIKWPLTVRFKLPN                      |
| TbPIF6 | 563 | -----GPALPVVCFATSG-----   |
|        |     | Motif V   |
| SCRRM3 | 621 | NKYIHELIMVPERFPIDIPRENUGLERTQIPLMLCMALSIHKAQGGOTIQRILKVDLRR-IFEAGQVYVALSRAVMTMDTLQVLNFDPG |
| SCPIF1 | 666 | MSTRMVLVEPEDWAIEDENEKPLVSRVOLPLMLAWSLSIHKSQGGOTLPKVKVDLRR-VFEKGOAYVALSRAVSREGLOVLNFDRT    |
| HSPIF1 | 520 | ---VTEVIHADRWTVQATGG-QLLSRQQLPLQLAWAMSIHKSQGGMTLDCEIISLGR-VFASGOAYVALSRAVSREGLOVLNFDPPM   |
| SPPFH1 | 683 | GGERTIVVQRETWNIELPNCGEVQASRSQIPLILAYASIHKAQGGOTLDRVKVDLGR-VFEKGOAYVALSRAVTOEGLQVLNFSPA    |
| TbPIF6 | 576 | G--EEVLVPRVSMEVLGPEGRVIAATRTQIPLQLSWALTVHRAQGGMTLPLVSVRLNKCFFDCGOAYVALSRAVSREDIMLTAFDPS   |
|        |     | Motif VI  |
| SCRRM3 | 705 | KTRTNERVKDFYKRRETLK-----  |
| SCPIF1 | 750 | RKKHQKVIDFYLTLSAESAYKQLEADEQVKKRKLIDYAPGPKYKAKSKSNSPAP-----ISATTQSNNGI                    |
| HSPIF1 | 600 | AVRCDPRVLHXYATIRRRGRSLSPESDDDEAASDQENMDPIL-----   |
| SPPFH1 | 767 | KVMAHPKVQFYKQLASVNGLPRI RNENKAPVQMRGVKNK-----   |
| TbPIF6 | 659 | AFADARAVAFYEKNFPAQRQSVEDTECELVITIKGKTRAKHPRSQGEKNSVDEGGNAPEEHPLRTDAAFTAYHDLDSQVSTDMPL     |
| SCRRM3 |     | -----   |
| SCPIF1 | 818 | AAMLQRHSRKRFFQLKKESNSNQVHSLVSDEPRGQDTEHILE-----   |
| HSPIF1 |     | -----   |
| SPPFH1 |     | -----   |
| TbPIF6 | 744 | VPQPPRKKRMLVEELPQVTSIPAIPNFTQESNNGDANSQLQHXFSQNKLMVDDDD                                   |

Figure 3-8 continued

Unsurprisingly, a phylogenetic tree generated from the alignment in Figure 3-8 (using CLC Genomics Workbench 6) shows *T. brucei*, *T. cruzi* and *L. major* PIF6 proteins to be much diverged from the other PIF1-like proteins analysed (Fig. 3-9), even from other kinetoplastid PIF1-like proteins in the alignment. This may be because it has been proposed that the expansion of the *T. brucei* Pif1 helicase family is due to requirements for kDNA maintenance, whereas TbPIF6 has been proposed to be nuclear (Liu *et al.*, 2009a). These analyses of TbPIF6 clearly demonstrate that TbPIF6 is a PIF1-like helicase, though highly diverged from other characterised eukaryotic, including kinetoplastid, PIF1-helicases analysed. Note this alignment excludes the large number of predicted PIF1-like proteins that were highly scoring hits in P-BLAST analysis.



**Figure 3-9 Neighbour-joining tree of eukaryotic Pif1 family helicases**  
A phylogenetic tree was constructed (using the neighbour joining method) from a multiple sequence alignment of alignment of Pif1 helicase family proteins (each constructed using CLC Genomics Workbench 6). TbPIF6 is highlighted in yellow. ARATH, *Arabidopsis thaliana*; Lm, *Leishmania major*; Tb, *T. brucei*; Tc, *Trypanosoma cruzi*; Sc, *S. cerevisiae*; Sp, *S.pombe*; Hs, *H. sapiens*. Accession numbers for sequences can be found in Appendix 7.2.

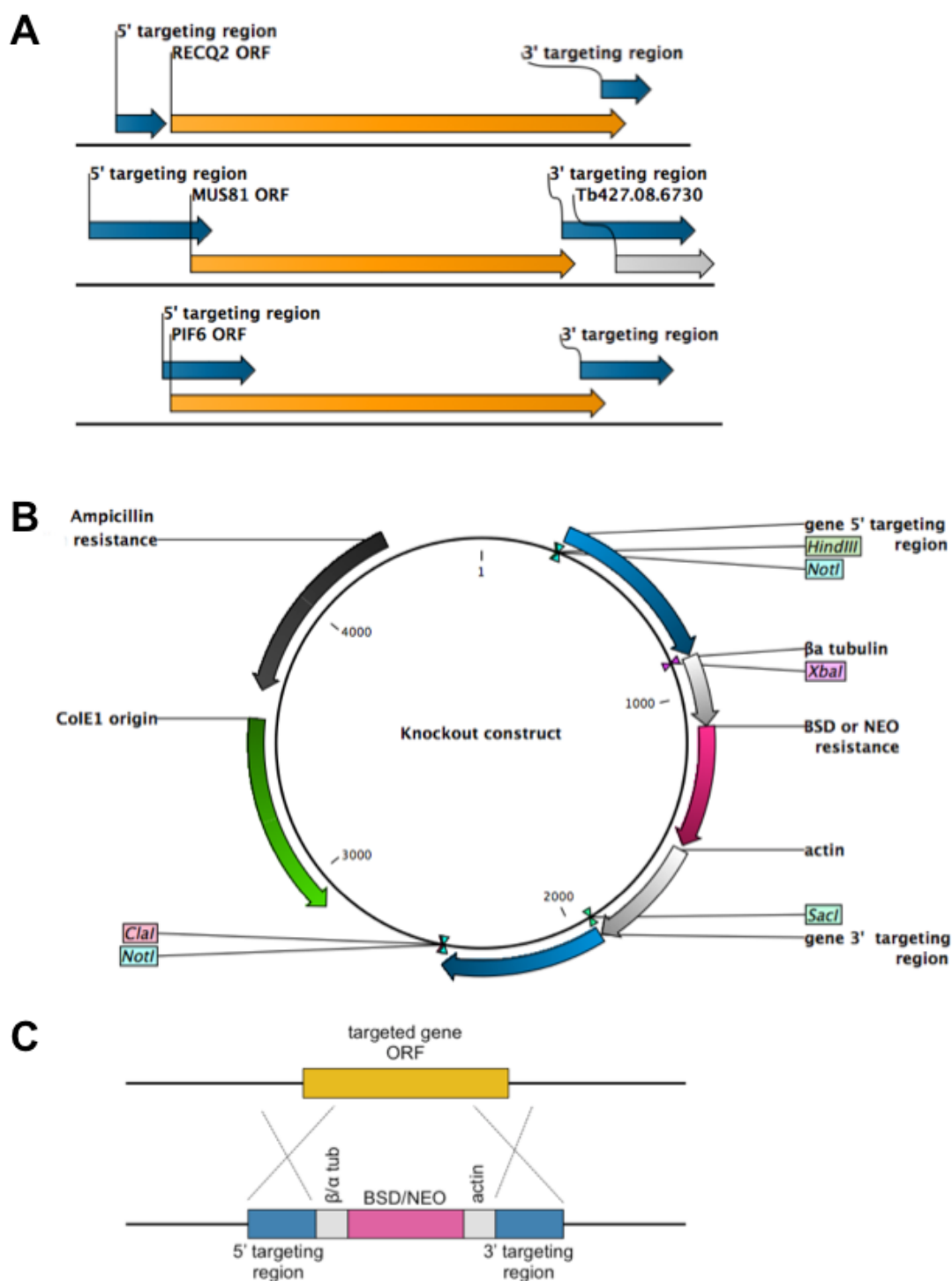
### 3.3 Generation of knockout mutants

#### 3.3.1 Knockout strategy

Heterozygous (+/-) and homozygous (-/-) knockout mutants for each gene were generated in Lister 427 *T. brucei* by deleting the entire ORF as shown in Figure

3-10A. All mutant cell lines were generated using the same vector backbone,  $\Delta ORF::BSD$  and  $\Delta ORF::NEO$  (BSD, blasticidin; NEO, neomycin), and examples of the knockout construct used are shown in Figure 3-10B (gift, Marko Prorocic). These constructs contain regions homologous to the 5' and 3' ends of the targeted gene flanking an antibiotic resistance cassette. These 5' and 3' regions were used as the targeting sequence to facilitate homologous recombination and replacement of the target ORF.

A ~500-800 bp region of the target gene's 5' ORF/5' UTR and a ~500-800 bp region of the target gene's 3' ORF/3' UTR were amplified by PCR and cloned into the knockout construct, flanking a *BSD* or *NEO* resistance cassette. The 5' UTR of the target gene was amplified by PCR using a forward primer that contained a *NotI* restriction site at the 5' end of the primer, followed by a *HindIII* restriction site and a reverse primer that contained an *XbaI* restriction site at the 5' end of the primer. The PCR amplified product was therefore 5'-*NotI*-*HindIII*-gene specific sequence-*XbaI*-3'. The 3' UTR of the target gene was amplified by PCR similarly, with the modification that the *NotI* restriction site was at the 3' end of the amplified product. The PCR amplified product was therefore 5'-*SacI*-gene specific sequence-*Clal*-*NotI*-3'. The regions PCR amplified for each construct are shown in Figure 3-10A, which in most cases included sequence from the target gene ORF as well as UTR sequence. The 3' targeting region PCR amplified in the *MUS81* knockout construct included the first 190 bp of the ORF of the gene immediately downstream of *MUS81* (Tb427.08.6730). PCR-specific details are listed in Figure 3-10. Integration of the knockout cassette however should not result in loss of or disruption to Tb427.08.6730 sequence.



**Figure 3-10 Schematic of knockout constructs and targeting strategy**  
 (A) PCR was used to amplify a 5' and a 3' region (blue arrows) of the target gene, in most cases including part of the ORF (orange arrow) and UTR of the target gene. These PCR products were the 'targeting regions' cloned into the knockout plasmid (see (b)). In the case of *MUS81*, the amplified 3' targeting region included the 3' 190 bp of the downstream gene (Tb427.08.6730). The primers used to generate the targeting regions were: *RECQ2*: 5' primers #30 & #31, 3' primers #32 & #33. *MUS81*: 5' primers #34 & #35, 3' primers #36 & #37. *PIF6*: 5' primers #38 & #39, 3' #40 & #41. Primer sequences can be found in Appendix 7.1.  
 (B) Generalised schematic of the knockout constructs. 3' and 5' targeting regions of the target gene (blue) flank the resistance cassette containing a *NEO* or *BSD* resistance gene (pink). *NotI* (as indicated) was used to linearise the constructs and restriction enzyme sites are indicated. Sizes shown (bp), length of the 5' and 3' targeting fragment lengths were specific to the gene targeted. (C) Knockout strategy used to generate heterozygous (+/-)

and knockout (-/-) mutants. Generalised representation of the linearised knockout *BSD* and *NEO* resistance constructs (bottom) relative to the targeted gene (ORF, orange) and its flanking regions (black line). The 5' and 3' targeting regions in the constructs correspond to the 5'-most or 3'-most regions of the targeted ORF and its flanking regions.  $\beta/\alpha$  tub,  $\beta/\alpha$  tubulin; actin IR, actin intergenic region; BSD, blasticidin resistance; NEO, neomycin resistance. Not to scale.

A generalised schematic of the knockout plasmids used is shown in Figure 3-10B, though the size of the targeting fragments differed depending on the gene targeted. To allow selection of successful *T. brucei* transformants, a BSD or NEO resistance cassette was cloned between the UTR flanks. In each construct the resistance gene is flanked upstream by a beta-alpha tubulin intergenic sequence and downstream by an actin intergenic sequence, to allow trans-splicing and polyadenylation respectively, and to generate antibiotic resistance ORF-containing mRNAs.

In a four-fragment ligation reaction the HindIII/XbaI digested 5' and ClaI/SacI digested 3' targeting fragment PCR products were ligated to the resistance cassette, which had been cut with XbaI (5') and SacI (3'), and to the plasmid backbone, which had been cut with ClaI and HindIII. This ligation produced a final plasmid that contained the resistance cassette, flanked 5' by the 5' region of the target gene and 3' by the 3' region of the target gene. This region, the knockout cassette, could be excised from the plasmid by digestion with NotI utilising the NotI restriction sites from the gene targeting PCR products. All plasmids were checked using restriction digest pattern analysis and sequencing. Approximately 10  $\mu$ g of plasmid DNA was digested and concentrated using a Zymogen Clean and Concentrate DNA concentrator kit prior to transformation, as described in Section 2.8.8.

Knockout constructs were linearised using NotI. Integration into the genome and replacement of the targeted gene was facilitated by the 5' and 3' targeting regions of the linearised knockout plasmid (Fig. 3-10C).

### 3.3.2 Generation of knockout mutants in the Lister 427 cell line

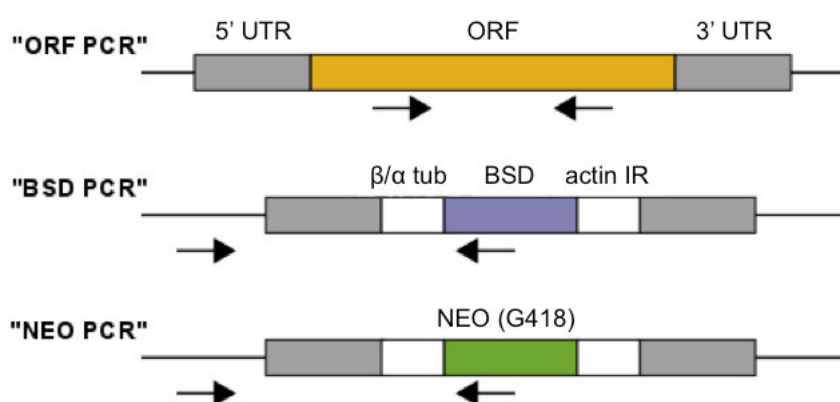
Wild type BSF Lister 427 cells were transformed (as described in Section 2.1.3) with  $\Delta RECQ2::BSD$ ,  $\Delta MUS81::BSD$  or  $\Delta PIF6::BSD$  constructs to attempt to generate *recq2*<sup>+/-</sup>, *mus81*<sup>+/-</sup> or *pif6*<sup>+/-</sup> cell lines respectively. In all cases, antibiotic resistant clones were selected using 10  $\mu$ g.mL<sup>-1</sup> blasticidin. However,

no surviving clones were obtained from the  $\Delta MUS81::BSD$  transformation, and so it was attempted to generate  $mus81^{+/-}$  cells by transforming the  $\Delta MUS81::NEO$  construct; selection with  $5 \mu\text{g.mL}^{-1}$  G418 allowed a number of surviving clones to be obtained. Sequencing of the  $\Delta MUS81::BSD$  construct revealed a non-synonymous mutation in the blasticidin resistance gene. The blasticidin resistance cassette  $\Delta MUS81::BSD$  was subsequently replaced and the plasmid sequenced to confirm the absence of mutations.

The generation of  $-/-$  mutants was confirmed by PCR analysis (see Section 3.1.3 below) performed on genomic DNA extracted from six different heterozygous clones for each gene, all of which contained the knockout construct integrated at the correct locus as expected. One confirmed  $+/-$  clone was transformed with the  $\Delta \text{gene of interest}::NEO$  construct (or blasticidin construct in the case of  $mus81^{+/-}$ ) and antibiotic resistant clones selected with  $5 \mu\text{g.mL}^{-1}$  blasticidin and  $2.5 \mu\text{g.mL}^{-1}$  neomycin.

### 3.3.3 Confirmation of knockout mutants by PCR

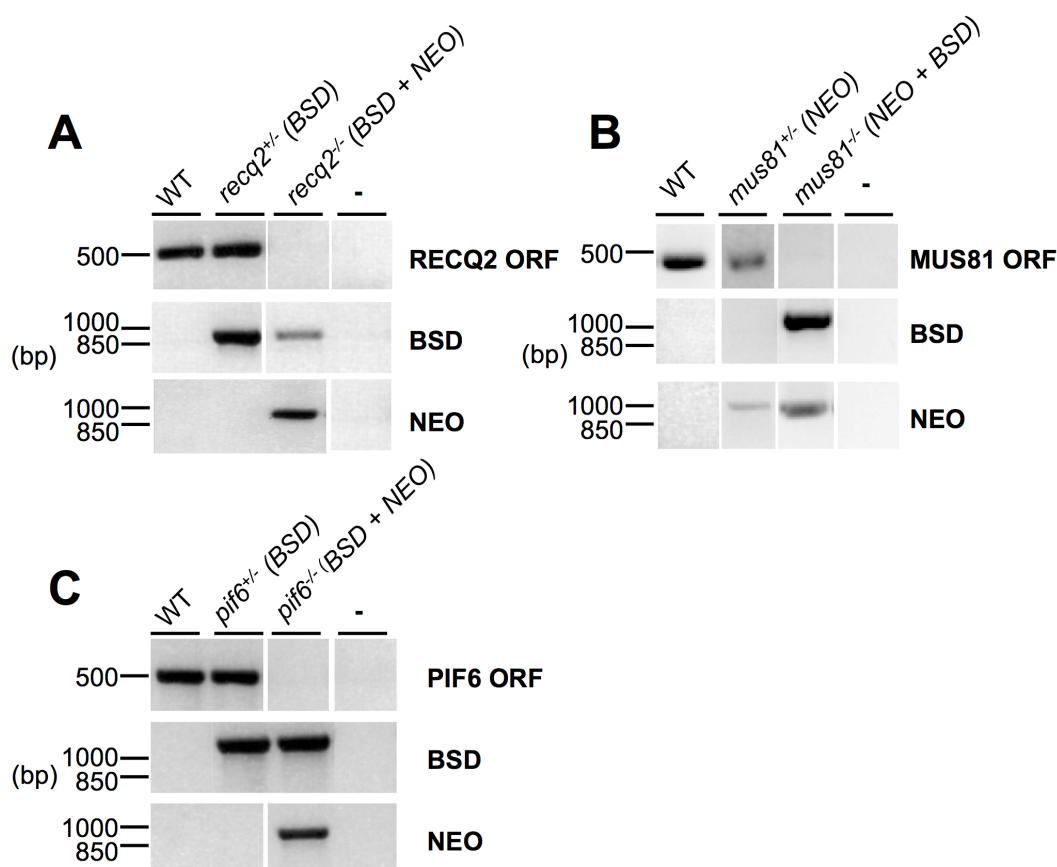
The generation of  $+/-$  and  $-/-$  mutants was confirmed by PCR performed on genomic DNA extracted from putative  $+/-$  (BSD; RECQ2 and PIF6),  $+/-$  (NEO; MUS81) and  $-/-$  (BSD+NEO) transformant clones. Three different PCRs were used, as shown in Figure 3-11, to confirm the correct integration of the BSD and NEO knockout cassettes relative to the target gene (“NEO” and “BSD” PCRs) and the presence of the targeted gene ORF.



**Figure 3-11** Diagnostic PCRs to confirm knockout of target genes  
A ~500 bp region of the open reading frame was amplified ("ORF PCR") and a ~1000 bp region was amplified using a forward primer lying upstream of the 5' UTR region present in the knockout construct and a reverse primer specific to the *BSD* or *NEO* gene ("BSD PCR" and "NEO PCR") respectively. RECQ2 "ORF PCR" primers #44 and #45, 498 bp; "BSD PCR"

primers #51 and #154, 1365 bp; “NEO PCR” primers #51 and #155, 1008 bp. MUS81”ORF PCR” primers #46 and #47, 494 bp; “BSD PCR” primers #52 and #154; 1331 bp, “NEO PCR” primers #52 and #155, 974 bp. PIF6 “ORF PCR” primers #48 and #49, 500 bp; “BSD PCR” primers #53 and #154, 1255 bp; “NEO PCR” primers #53 and #155, 898 bp. Sequences of all primers are in Appendix 7.1. Black arrows indicate primer binding sites. Not to scale.

The agarose gel in Figure 3-12 shows confirmation, using the above PCRs, of the integration of the knockout cassettes in each cell line and the removal of the cognate ORFs after the second round of transformation. Only one knockout cassette could be PCR amplified (“BSD/NEO PCR”) in each heterozygote line, whereas both cassettes could be PCR amplified in the double knockout cell lines. Furthermore, the ORF of the targeted gene could no longer be PCR amplified in double knockout cell lines, indicating successful knockout of each gene.



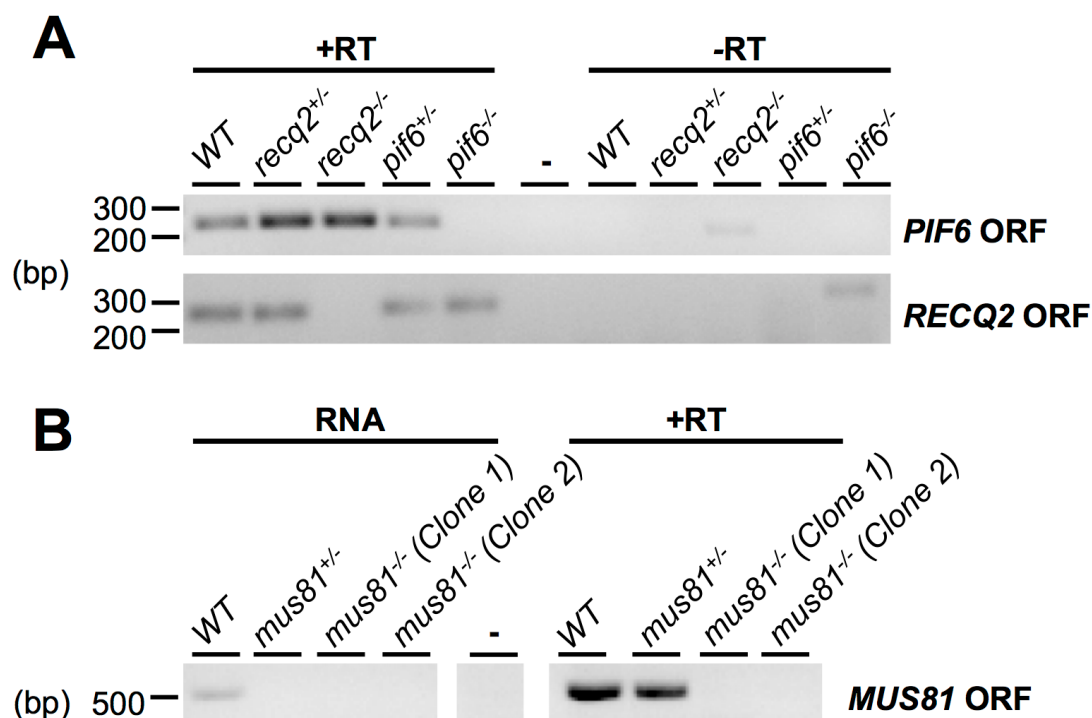
**Figure 3-12** Confirmation by PCR of knockout of RECQ2, MUS81 and PIF6. Agarose gels of PCR products generated from genomic extracted from DNA wild type and *recq2*, *mus81* and *pif6* mutant cells. A 500 bp region of the open reading frame was amplified (“ORF PCR”) and a ~1000 bp region was amplified using a forward primer lying upstream of the 5’ UTR region present in the knockout construct and a reverse primer specific to the *BSD* or *NEO* gene (“BSD PCR” and “NEO PCR”) respectively. These primers are described in Figure 3-11 and their sequences can be found in Appendix 7.1

### 3.3.4 Confirmation of knockout mutants by RT-PCR

Reverse transcription PCR (RT-PCR) was used to confirm absence of expression of the knocked-out gene in each  $-/-$  cell line. RNA extracted from wild type,  $+/-$  and  $-/-$  cell lines was reverse transcribed (as described in sections 2.4.2 and 2.8.3). During reverse transcription, a no reverse transcriptase ( $-RT$ ) control was performed to assay for complete genomic DNA digestion. A region of the ORF of the deleted gene was amplified from the cDNA template, as well as a region of the ORF of another gene (*RECQ2* or *pif6* $^{-/-}$  mutants or *PIF6* for *recq2* $^{-/-}$  mutants) that was not targeted, as a positive control to confirm the presence of cDNA in the sample. This control PCR was not performed for *MUS81*, therefore the quality of the cDNA in *mus81* $^{-/-}$  preparations was uncertain. Sequences of all primers can be found in Appendix 7.1 and gene-specific PCR details are described in the legend of Figure 3-13.

The agarose gel in Figure 3-13 shows that the segment of the ORF of *RECQ2*, *MUS81* and *PIF6* could not be PCR amplified in their respective knockout cell lines, but that PCR amplification of the control gene was successful in the case of *recq2* and *pif6* mutants. In contrast, the expected PCR products were generated in the different  $+/-$  cells and in the untransformed wild type cells. There was faint PCR amplification in some  $-RT$  control samples, indicating DNA contamination in those RNA preparations. However, the ORF PCR for the targeted gene failed in knockout cell lines in both the  $+RT$  and  $-RT$  samples, confirming that knockout of the targeted genes was successful.





**Figure 3-13** RT-PCR confirmation of *recq2*, *mus81* and *pif6* mutants  
cDNA was synthesised from RNA extracted from wild type, *recq2*, *mus81* and *pif6* mutant cells. (A) A 245 bp region of the PIF6 ORF (primers #81/#82) and 232 bp region of the RECQ2 ORF (primers #77/#78) was PCR amplified from cDNA of wild type cells, *recq2* mutants and *pif6* mutants (+RT), as well as from samples in which no reverse transcriptase had been added (-RT). Distilled water was used as a negative control (-). (B) A 494 bp region of the MUS81 ORF (primers #46/#47) was PCR amplified from cDNA of wild type cells and *mus81* mutants (+RT), as well as from RNA used to synthesise the cDNA (RNA). Distilled water was used as a negative control. Dashes lined indicate different images aligned in this figure. Sizes shown, ladder (bp).

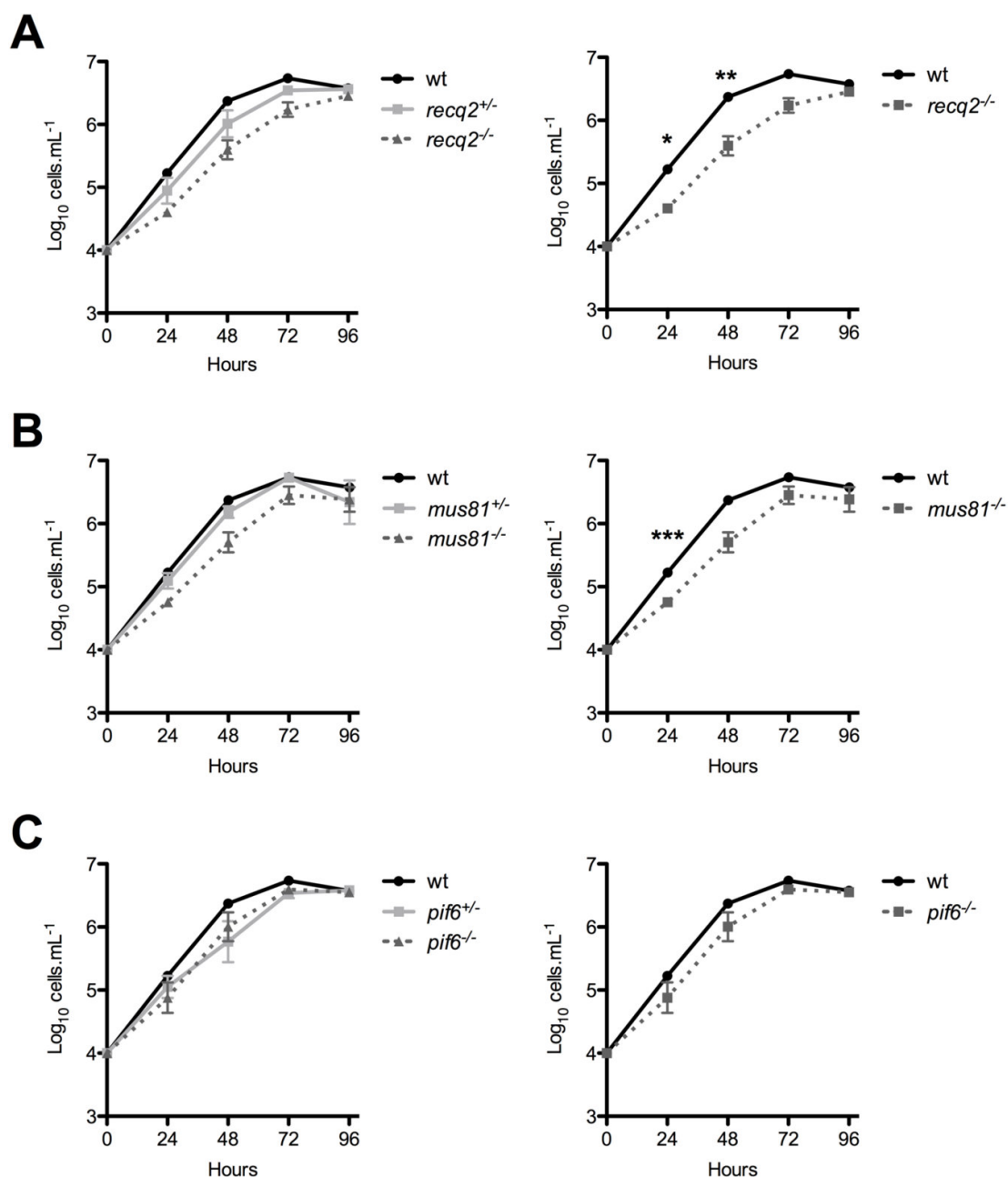
### 3.4 Phenotypic analysis of knockout mutants

#### 3.4.1 Analysis of in vitro growth

In order to understand if any of the gene knockouts had an effect on cell survival, the *in vitro* growth of *recq2*, *pif6* and *mus81* <sup>+/-</sup> and <sup>-/-</sup> mutants was compared to wild type cells by incubation of 1 mL cultures at a density of  $1 \times 10^4$  cells.mL<sup>-1</sup> in the absence of selective drugs and the density counted every 24 hours up to a maximum of 96 hours using a haemocytometer. Cultures were not diluted during the growth curve. All growth curves were repeated three times and the average and standard error of the mean (SEM) of these counts were plotted logarithmically (Fig. 3-14).

A growth defect was observed for *recq2*<sup>-/-</sup> mutants compared to wild type cells (Fig. 3-14A), which was statistically significant at 24 and 48 hours ( $p=0.018$  and  $p=0.037$  respectively, using a 2 sample t-test). In addition, there appeared to be

a minor effect on growth in the *req2*<sup>+/-</sup> mutants, perhaps suggesting some impediment after loss of one allele. A growth defect similar to the *recq2*<sup>-/-</sup> mutants was also observed in *mus81*<sup>-/-</sup> mutants (Fig. 3-14B), though this was only significant at the 24 hour time point (p=0.005 using a 2 sample t-test). For these mutant series there was, however, no obvious growth impediment in the *mus81*<sup>+/-</sup> cells, suggesting any effects are only manifest after both alleles are deleted. No defect in growth was found to be statistically significant for either *pif6* mutant (Fig. 3-14C), though the growth curves suggest a minor growth impediment that is of the same magnitude in both the +/- and -/- cells. Interestingly, this appears to correlate with several further assays (below), which suggest that loss of one *pif6* allele has phenotypic consequences that are not increased further after loss of the second allele.



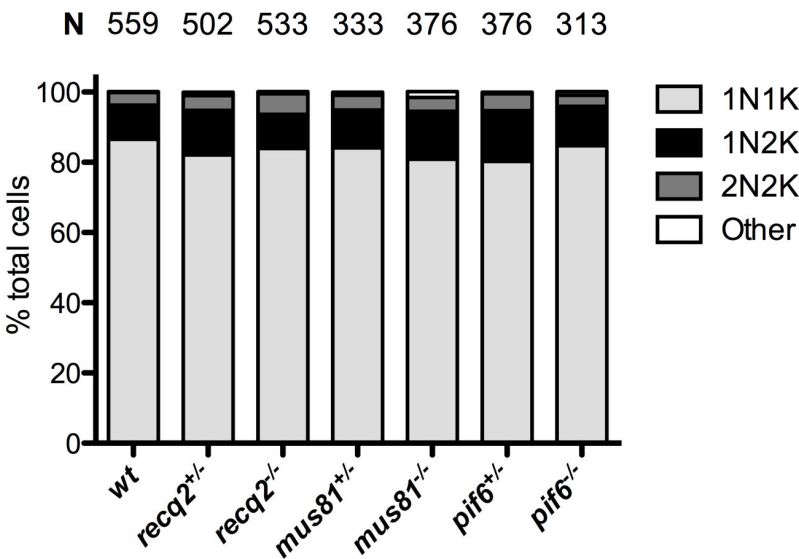
**Figure 3-14** *In vitro* growth analysis of *recq2*, *mus81* and *pif6* mutants. The cell density of wild type and *recq2* mutants (A), *mus81* mutants (B) and *pif6* mutants (C). Cultures were counted every 24 hours up to a maximum of 96 hours using a haemocytometer starting from a cell density of  $1 \times 10^4$  cells.mL<sup>-1</sup>. Growth curves were repeated three times and the mean cell density and standard error of the mean (SEM) are plotted logarithmically. For clarity, *in vitro* growth of wild type, +/- and -/- mutants is plotted on the left and the growth of wild type and -/- cells only is plotted on the right. Significant differences between the means are indicated by \* and were calculated using Student's t-test; \*  $p = 0.018$ ; \*\*  $p = 0.037$ ; \*\*\*  $p = 0.005$ .

### 3.4.2 Cell cycle analysis

To investigate if the growth impediments (above) can be explained by delay in completion of the cell cycle, it was examined whether there were detectable cell cycle alterations in *recq2*, *mus81* and *pif6* mutants. For this purpose, cells

were fixed, stained with DAPI and scored by cell cycle stage. Cells in the G1 and S phases of the cell cycle contain one nucleus and one kinetoplast (1N1K). Kinetoplast division occurs before nuclear division and thus cells in the G2 phase contain one nucleus and two kinetoplasts (1N2K) and cells with two nuclei and two kinetoplasts are in M phase (2N2K). Following M phase, cytokinesis results in two 1N1K cells (Woodward & Gull, 1990). The expected distribution of detectable cell cycle phases in wild type blood stream form cells is: ~80% 1N1K, ~10% 1N2K, ~5% 2N2K (Benmerzouga *et al.*, 2013).

Figure 3-15 shows cell cycle analysis of *mus81*, *pif6* and *recq2* mutants. The distribution of cell cycle stages in all of the +/- or -/- cell lines was similar to the expected distribution, indicating no detectable defects in the cell cycle using this method. Therefore, although *in vitro* growth analysis showed that at least the *recq2*<sup>-/-</sup> and *mus81*<sup>-/-</sup> mutants have a growth defect, this is not accompanied by changes in progression through the cell cycle.



**Figure 3-15 Cell cycle analysis of *recq2*, *mus81* and *pif6* mutants**  
 Logarithmically growing wild type and mutant cells were fixed and stained with DAPI to visualise the nucleus and kinetoplast. Cells in captured images were counted (using Fuji ImageJ) and categorised as 1N1K, 1N2K, 2N2K or 'other' for cells that did not conform to these categories. N, nucleus, K, kinetoplast. The number of cells counted is indicated above each bar (N).

### 3.4.3 Analysis of DNA damage sensitivity

The three genes under investigation have each been shown to act in genome repair in other eukaryotes. To ask if they also act in similar pathways of repair

in *T. brucei*, analysis of the growth and clonal survival of cells following exogenous DNA damage was performed, assessing the DNA damage sensitivity of the mutants. Three DNA damaging agents with different modes of action and resultant damage were used: hydroxyurea (HU), methyl methanesulfonate (MMS) and phleomycin. HU depletes the cellular dNTP pool (Bianchi *et al.*, 1986), resulting in replication stalling. Stalled replication forks can subsequently collapse, resulting in DNA breaks. MMS is an alkylating agent, which methylates purines at the 7' position of guanosine residues and the 3' position of adenine residues (Brookes & Lawley, 1961). MMS causes DNA breaks, though it is not clear whether it produces these directly or whether DNA breaks are formed through activities resulting from DNA alkylation (Lundin *et al.*, 2005; Wyatt & Pittman, 2006). MMS also perturbs replication, slowing the progression of the replication fork due to alkylated nucleotides physically blocking replication fork elongation (Groth *et al.*, 2010; Tercero & Diffley, 2001). Phleomycin blocks the activity of DNA polymerase, inhibiting DNA synthesis and resulting in formation of single and double strand DNA breaks, primarily the latter (Falaschi & Kornberg, 1964; Reiter *et al.*, 1972).

The sensitivity of *recq2*, *pif6* and *mus81* mutants to these damaging agents was assayed using clonal survival assays and measuring the outgrowth of a single cell in the presence of a damaging agent. Prior to this and in order to determine the appropriate concentrations of each damaging agent, *in vitro* growth was assayed by growth curve analysis in the presence of a damaging agent.

#### **3.4.3.1 Analysis of *in vitro* growth after DNA damage**

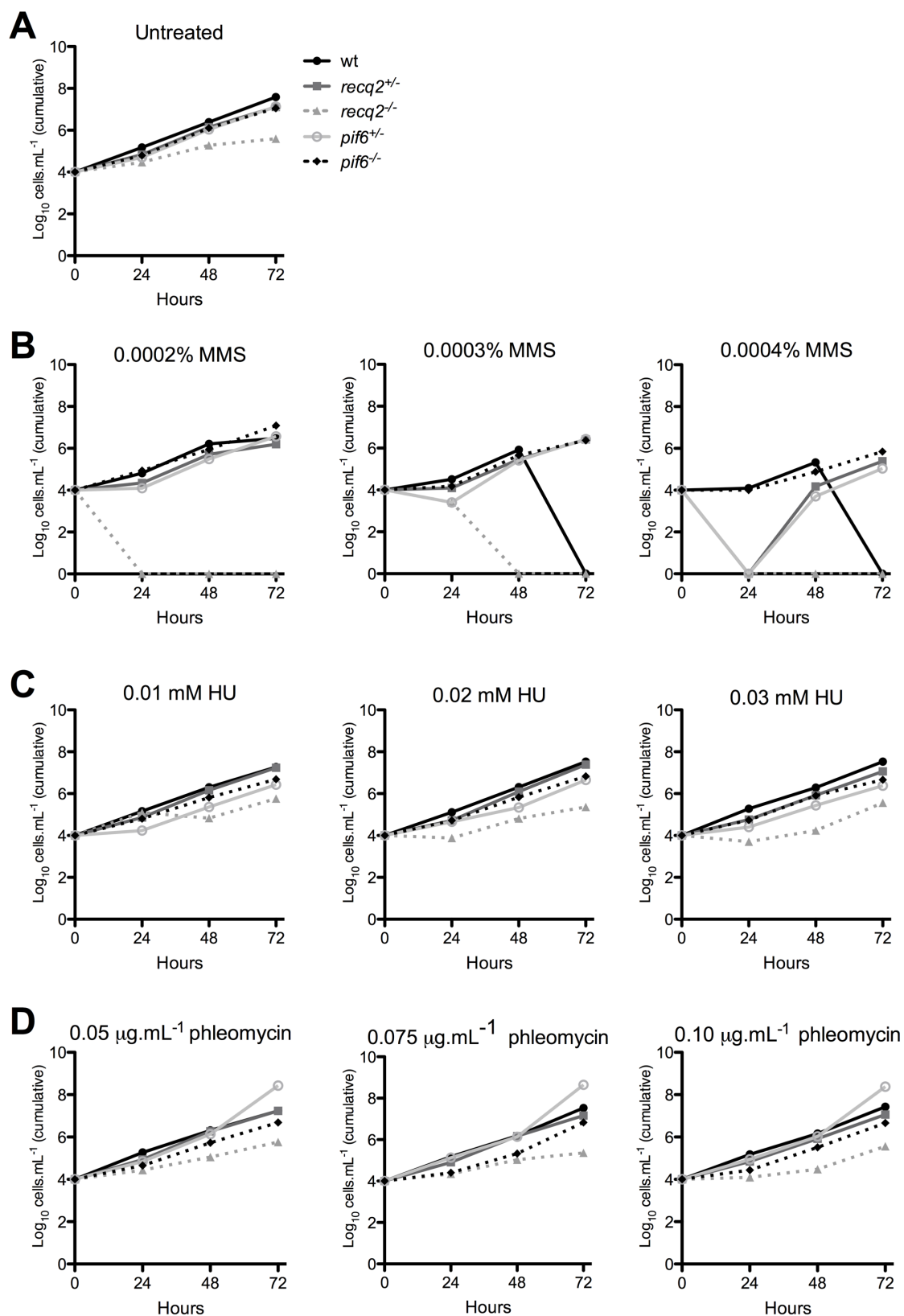
To determine the appropriate concentrations to use in subsequent clonal survival assays, growth curve analysis of *pif6* and *recq2* mutants was performed. The same analysis was not performed for *mus81* mutants.

The initial concentrations of MMS chosen for *in vitro* growth analysis were 0.0002%, 0.0003% and 0.0004%, based on our lab's previous experience. The growth of untreated cells is shown in Figure 3-16A for comparison. Following MMS treatment, wild type cells exhibited slower growth as expected, which was detectable at 0.0002% and became more pronounced at 0.0003% and 0.0004%. The very slow growth of *recq2*<sup>-/-</sup> mutants displayed an increased impairment in

growth relative to wild type cells and *recq2*<sup>+/-</sup> mutants at all concentrations of MMS tested (Fig. 3-16B), suggesting that null mutants are more sensitive to MMS-induced DNA damage. No cells were counted in 0.0003% and 0.0004% MMS treated wild type cultures at 72 hours. This is most likely due to technical error in passage of cultures at 48 hours. The very slow growth of *recq2*<sup>-/-</sup> at even the lowest concentration of MMS (0.0002%) suggested that an additional, lower concentration (0.0001%) of MMS should be used in subsequent clonal survival assays (see below). The growth of *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> cells was similar to wild type cells after MMS treatment (Fig. 3-16B). These data suggested that *pif6* mutants have no difference in MMS-induced damage sensitivity or alternatively, any phenotype is not detectable in this assay.

The HU concentrations used in the analysis of *in vitro* growth were 0.01 mM, 0.02 mM and 0.03 mM, based on Kim & Cross (2010). Following HU treatment, no obvious effect on growth was observed for wild type cells (Fig. 3-16C). *recq2*<sup>+/-</sup> mutants grew similarly to wild type cells under HU treatment but *recq2*<sup>-/-</sup> mutants appeared to grow more slowly in the presence of HU. In addition, there was a minor growth defect in both *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> mutants. These data suggested that the HU concentration range should be increased to attempt to observe a more obvious phenotype in *pif6* mutants.

Little effect on the growth of wild type or *recq2*<sup>+/-</sup> cells was observed during growth in the presence of phleomycin (Fig. 3-16D). However, the growth of phleomycin treated *recq2*<sup>-/-</sup> cells was slower compared to untreated *recq2*<sup>-/-</sup> cells and wild type phleomycin treated cells. Interestingly, phleomycin-treated *pif6*<sup>+/-</sup> cells appeared to grow faster than untreated cells, although this effect was only observed at 72 hours. In contrast, no effect on growth was observed for *pif6*<sup>-/-</sup> phleomycin-treated cells. The absence of an observable phleomycin treated growth phenotype in most of the cell lines suggested that the phleomycin concentration range should be increased in subsequent clonal survival assays.



**Figure 3-16** *In vitro* growth of *recq2* and *pif6* mutants in the presence of DNA damaging agents

Wild type, *recq2* mutant and *pif6* mutant cells at a starting density of  $1 \times 10^4$  cells.mL<sup>-1</sup> were incubated for 72 hours in the presence of (A) no damaging agents, (B) varying concentrations of MMS, (C) varying concentrations of hydroxyurea (HU) or (D) varying concentrations of phleomycin. The cell density was counted every 24 hours using a haemocytometer. Cumulative growth is plotted logarithmically.

### 3.4.3.2 Analysis of clonal survival after DNA damage

In order to analyse the clonal survival of *recq2*, *mus81* and *pif6* mutants, treated and untreated cultures were diluted to one cell per well and the number of wells on a 96 well plate with live cells was counted after ~10 days incubation (see Section 2.2.3.2 for more details).

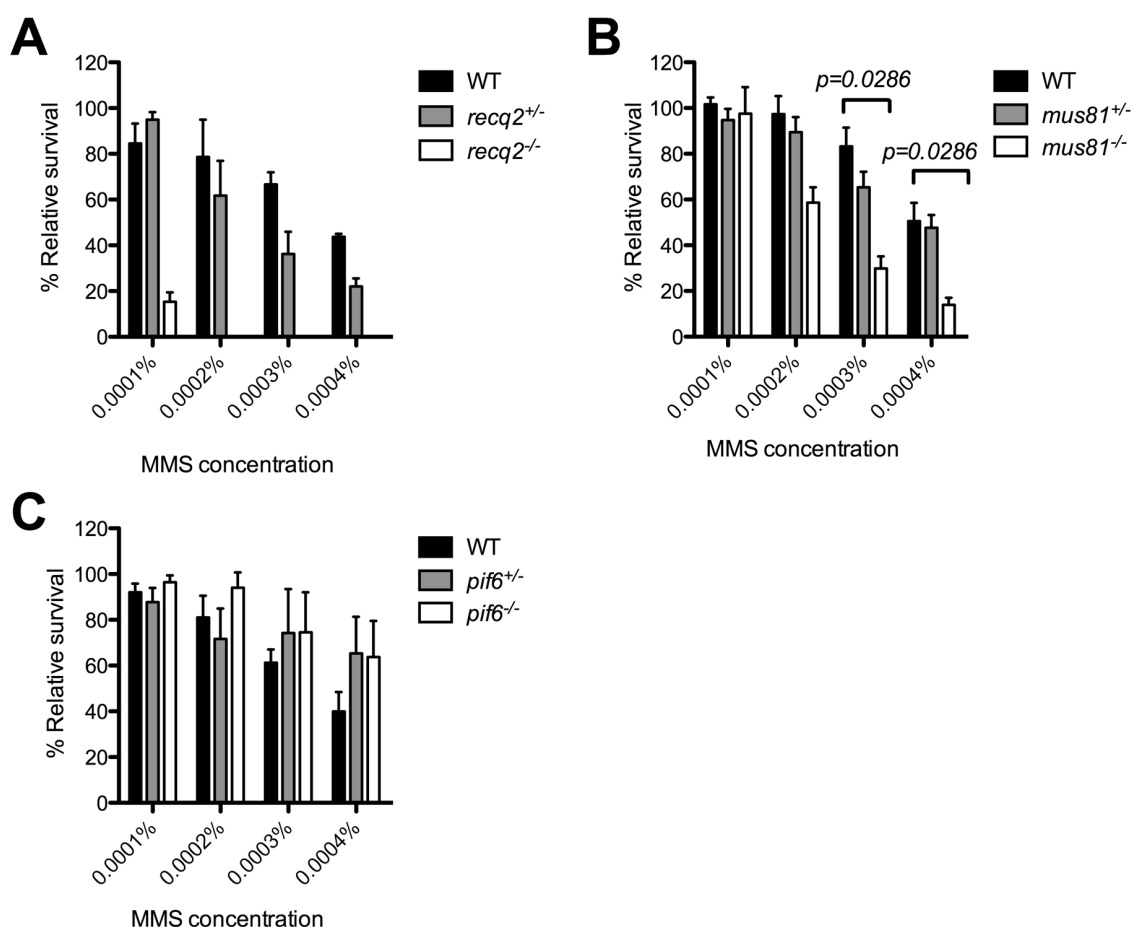
MMS concentrations used (0.0001%, 0.0002%, 0.0003% and 0.0004%) were determined from previous *in vitro* growth curve analysis (Section 3.4.3.1). The survival of wild type cultures decreased with increasing MMS concentration (Fig. 3-17). At MMS concentrations at and above 0.0002%, *recq2*<sup>+/-</sup> survival was lower than wild type and *recq2*<sup>-/-</sup> cells displayed a severe sensitivity to MMS, though not statistically significant using a Mann-Whitney U test. These data indicate that *recq2*<sup>-/-</sup> mutants are more sensitive to MMS and that mild MMS sensitivity is caused by loss of even one allele, indicating that RECQ2 is involved in the *T. brucei* response to MMS-induced damage. MMS sensitivity has previously been reported in RecQ helicase mutants. *S. cerevisiae* SGS1 mutants (Mullen *et al.*, 2000) as well as human *blm*<sup>-/-</sup> (Wang *et al.*, 2000) and chicken DT40 cells (Imamura *et al.*, 2001) are hypersensitive to MMS.

No change in MMS sensitivity was seen in *mus81*<sup>+/-</sup> cells relative to wild type (Fig. 3-17). At higher concentrations of MMS (0.0003% and 0.0004% MMS), *mus81*<sup>-/-</sup> displayed approximately three-fold reduced survival relative to wild type using a Mann-Whitney U test (0.0003% MMS: P=0.0286, WT n=4 median=79.05, *mus81*<sup>-/-</sup> n=4 median=33.05; 0.0004%: P=0.0286, WT n=4 median=47.65, *mus81*<sup>-/-</sup> n=4 median=13.25). These survival data indicate that MUS81 also acts in the response to MMS damage, though the reduced sensitivity of *mus81*<sup>-/-</sup> mutants relative to *recq2*<sup>-/-</sup> mutants suggests a greater role for the latter factor.

Following incubation with MMS, the survival of both *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> mutants appeared higher relative to wild type cultures, a quite different effect of these mutations to that of *mus81* and *recq2*. It is also notable that the effect of a single *pif6* allele knockout appeared the same as removing both alleles, consistent with the equivalent minor growth defects (Figure 3-14C). Despite this trend, there was considerable variation in survival of the *pif6* mutants in each individual experiment performed (Fig. 3-17) and there was no statistical



difference detected (Mann-Whitney U test) between wild type cells and *pif6* mutants. In fact, it was noted that assays conducted using *pif6* mutants that had been in culture for a longer period of time showed an increased MMS sensitivity than when *pif6* mutant cell lines were used after only a short period in culture (Appendix 7.2, Fig. 7-1). The basis for this variation is unclear, though it is possible that loss of PIF6 results in the selection for unknown, secondary mutations. Irrespective, it is clear that PIF6 does not play the same role in response to MMS damage as either RECQ2 or MUS81, and its absence may be beneficial in this circumstance.



**Figure 3-17** Clonal survival of *recq2*, *mus81* and *pif6* mutants in the presence of MMS. Wild type and *recq2* (A), *mus81* (B), and *pif6* (C) mutants were diluted to the equivalent of one cell per well over three 96 well plates in the presence of varying concentrations of MMS. Plates were incubated for 10 days or until wells were clearly positive or negative by colour (saturated cultures are indicated by yellow media). Positive wells were counted and the survival of MMS-treated cultures was calculated relative to untreated. Assays were repeated three times for *recq2* mutants, four times for *mus81* mutants and three times for *pif6* mutants. Mean survival (%) relative to untreated cells is plotted; bars represent standard error of the mean (SEM). Significant differences between the relative survival of different cell lines are indicated (Mann-Whitney U test - 0.0003% MMS: P=0.0286, WT n=4, median=79.05; *mus81*<sup>-/-</sup> n=4, median=33.05; 0.0004%: P=0.0286, WT n=4 median=47.65, *mus81*<sup>-/-</sup> n=4 median=13.25).

Initial clonal survival experiments were conducted using phleomycin at  $0.1 \mu\text{g.mL}^{-1}$ ,  $0.2 \mu\text{g.mL}^{-1}$  and  $0.3 \mu\text{g.mL}^{-1}$ , estimated from the lack of strong growth effects in the growth curve analysis (Section 3.4.3.1). However, these concentrations were found to be too high, as no cell line survived at concentrations higher than  $0.1 \mu\text{g.mL}^{-1}$  phleomycin (data not shown). The concentrations were thus lowered to  $0.05 \mu\text{g.mL}^{-1}$ ,  $0.075 \mu\text{g.mL}^{-1}$  and  $0.10 \mu\text{g.mL}^{-1}$ , the same concentrations used in the growth curve.

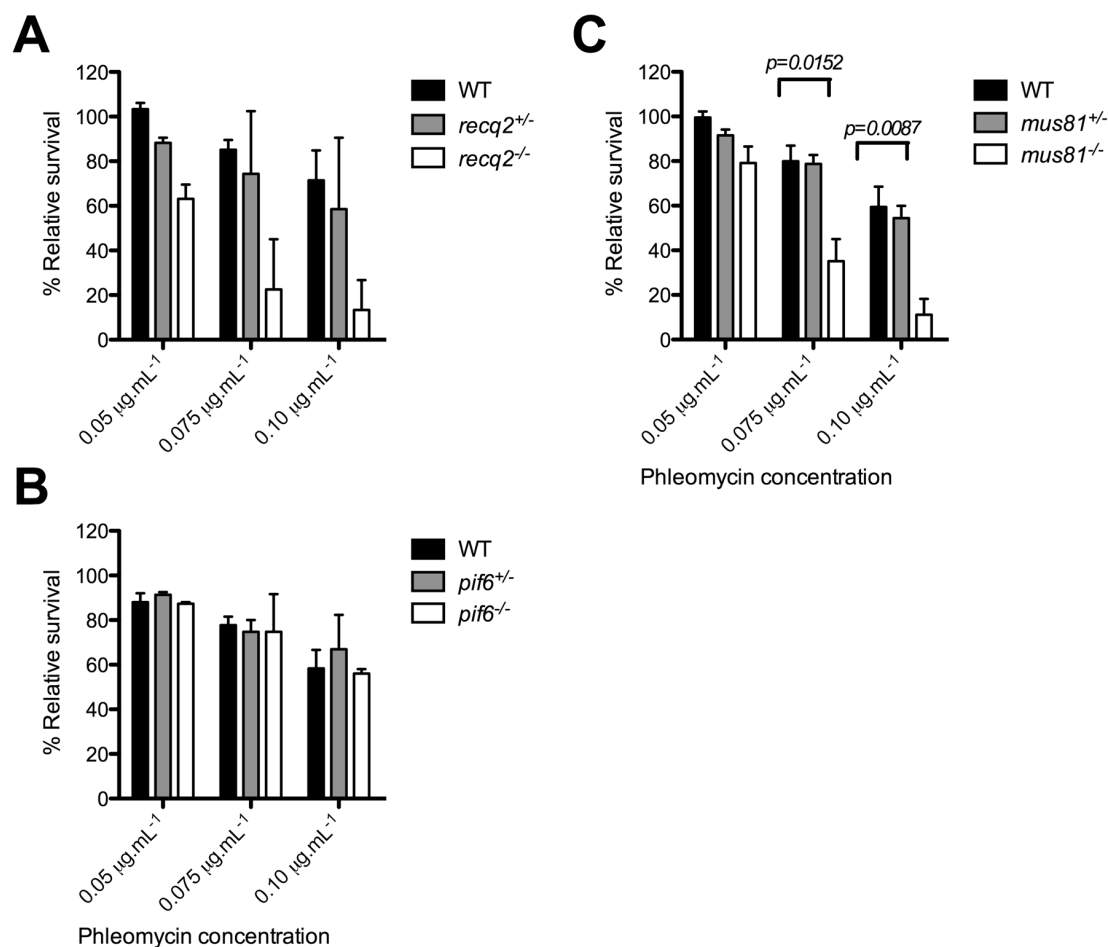
The clonal survival of *recq2* mutants in the presence of phleomycin treatment was assayed twice and in each of these experiments, *recq2*<sup>-/-</sup> mutants displayed reduced survival relative to wild type cells (Fig. 3-18), suggesting a role for RECQ2 in responding to phleomycin-induced damage. However, there was a high variation in the survival rate in the two repeats.

Survival of *mus81*<sup>+/-</sup> cells was similar to wild type at all concentrations of phleomycin treatment (Fig. 3-18). However, at  $0.075 \mu\text{g.mL}^{-1}$  and  $0.10 \mu\text{g.mL}^{-1}$  phleomycin *mus81*<sup>-/-</sup> survival was significantly lower than wild type using a Mann-Whitney U test ( $0.075 \mu\text{g.mL}^{-1}$ :  $P=0.0152$ , WT  $n=6$  median=86.2, *mus81*<sup>-/-</sup>  $n=6$  median=29.9;  $0.10 \mu\text{g.mL}^{-1}$ :  $P=0.0087$ , WT  $n=6$  median=62.36, *mus81*<sup>-/-</sup>  $n=6$  median=1.10). These data demonstrate that *mus81*<sup>-/-</sup>, but not *mus81*<sup>+/-</sup>, mutants have an increased sensitivity to phleomycin-induced damage and indicate that MUS81 is involved in the response to phleomycin damage.

No difference in sensitivity to phleomycin was observed in either *pif6* mutant compared to wild type cells (Fig. 3-18), suggesting PIF6 does not play a role in the response to phleomycin-induced damage and again indicating a distinct role to that of RECQ2 and MUS81.

Hydroxyurea concentrations used (0.02 mM, 0.03 mM and 0.04 mM) were determined from previous *in vitro* growth curve analysis. No effect on survival was observed after HU treatment of either wild type or *recq2*<sup>+/-</sup> cells (Fig. 3-19A). In contrast, *recq2*<sup>-/-</sup> cells displayed reduced survival compared to both cells at all concentrations of HU, with survival of *recq2*<sup>-/-</sup> at 0.04 mM approximately five-fold lower than wild type cells. These data suggest that RECQ2 acts in the *T. brucei* BSF cell response to HU damage.

Sensitivity of *mus81* mutants to HU was assayed twice and when data from the two assays were combined (Fig. 3-19B) no obvious change in sensitivity to HU was observed in *mus81* mutants. When these two assays are examined separately (Fig. 3-19D), *mus81*<sup>-/-</sup> cells displayed ~20% reduced survival compared to wild type at the highest HU concentration tested (0.04 mM) in only one of the experiments. These data suggest that MUS81 does not play a role in the response to HU, or at least has a very minor role.

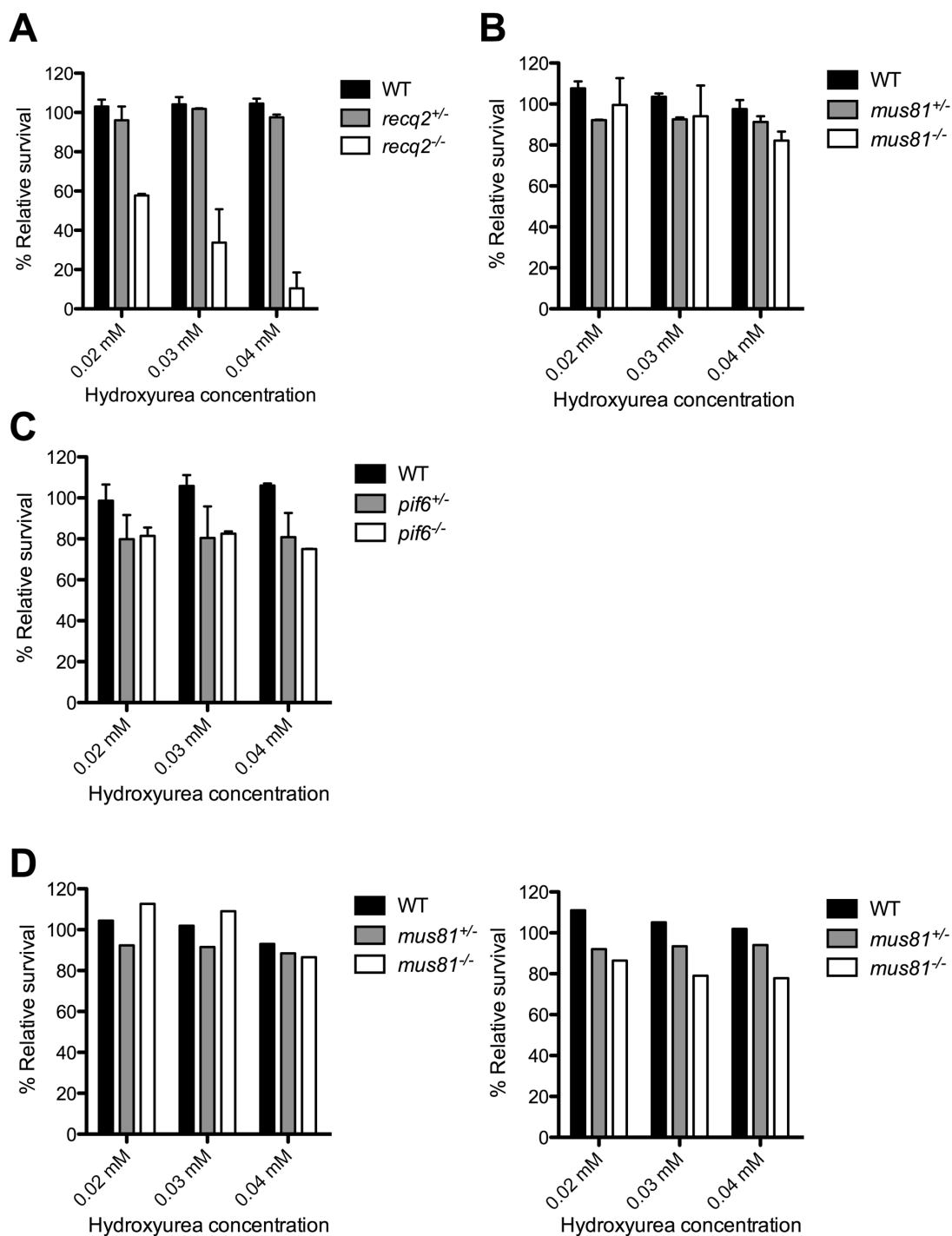


**Figure 3-18** Clonal survival of *recq2*, *mus81* and *pif6* mutants following treatment with phleomycin

Wild type and *recq2* (A), *mus81* (B), and *pif6* (C) mutants were diluted to the equivalent of one cell per well over three 96 well plates in the presence of varying concentrations of phleomycin. Plates were incubated for 10 days or until wells were clearly positive or negative by colour (saturated cultures are indicated by yellow media). Positive wells were counted and the survival of phleomycin-treated cultures was calculated relative to untreated. Assays were repeated twice for *recq2* mutants, six times for *mus81* mutants and twice for *pif6* mutants. Mean survival (%) relative to untreated cells is plotted; bars represent standard error of the mean (SEM). Significant differences between the relative survival of different cell lines are indicated (Mann-Whitney U test 0.075 µg.mL<sup>-1</sup>: *P*=0.0152, WT *n*=6 median=86.2, *mus81*<sup>-/-</sup> *n*=6 median=29.9; 0.10 µg.mL<sup>-1</sup>: *P*=0.0087, WT *n*=6 median=62.36, *mus81*<sup>-/-</sup> *n*=6 median=1.10).

A mild increase in HU sensitivity was observed in *pif6* mutants (Fig. 3-19C), survival of which was approximately 15% lower than wild type at 0.04 mM HU. Notably, and in common with the MMS data (Figure 3-17), both the *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> mutants displayed the same effect. Although these data are from only two repeats of this assay, and are perhaps surprising in that sensitivity increases with HU concentration (in contrast to *recq2*<sup>-/-</sup>), they suggest that PIF6 contributes to the response to HU. This contrasts with the lack of a detectable repair role against either MMS or phleomycin damage.

Taken together these data show that each of the knockout mutants displays a different DNA damage sensitivity phenotype. *recq2*<sup>-/-</sup> mutants displayed pronounced increased sensitivity to the DNA damaging agents MMS and hydroxyurea, and some evidence for increased sensitivity to phleomycin, suggesting RECQ2 acts in a broad range of repair roles. Though not as dramatic as in *recq2*<sup>-/-</sup> mutants, *mus81*<sup>-/-</sup> mutants also display an increased sensitivity to MMS compared to wild type and *mus81*<sup>+/-</sup> cells. *mus81*<sup>-/-</sup> cells were more sensitive than wild type or *mus81*<sup>+/-</sup> cells to phleomycin and, in this case, the sensitivity was similar to that of *recq2*<sup>-/-</sup> cells. Only weak evidence was found for a change in sensitivity to HU in *mus81*<sup>-/-</sup> cells. Thus, in total, MUS81 appears to display a similar broad range of repair activities to RECQ2, but for two forms of damage this role is substantially weaker and only for phleomycin damage does the putative endonuclease appear to compare in importance. In contrast to both *mus81*<sup>-/-</sup> and *recq2*<sup>-/-</sup> mutants, *pif6* mutants showed no increase in MMS or phleomycin sensitivity compared to wild type cells (indeed, if anything, *pif6* mutants displayed improved survival in response to MMS); instead, *pif6* mutant cells exhibited an increased sensitivity only to HU. Moreover, *pif6* mutants were notable for displaying no discernible difference in phenotypes between the +/- and -/- mutation configuration. PIF6 appears to provide a narrower DNA repair role in *T. brucei* than MUS81 or RECQ2.



**Figure 3-19** Clonal survival of *recq2*, *mus81* and *pif6* mutants in the presence of hydroxyurea

Wild type and *recq2* (A), *mus81* (B), and *pif6* (C) mutants were diluted to the equivalent of one cell per well over three 96 well plates, in the presence of varying concentrations of hydroxyurea (HU). Plates were incubated for 10 days or until wells were clearly positive or negative by colour (saturated cultures are indicated by yellow media). Positive wells were counted and the survival of HU-treated cultures was calculated relative to untreated. Assays were repeated twice. The two individual repeats assaying *mus81* mutants are shown in (D). Mean survival (%) relative to untreated cells is plotted; bars represent standard error of the mean (SEM).

## 3.5 Generation of re-expresser cell lines

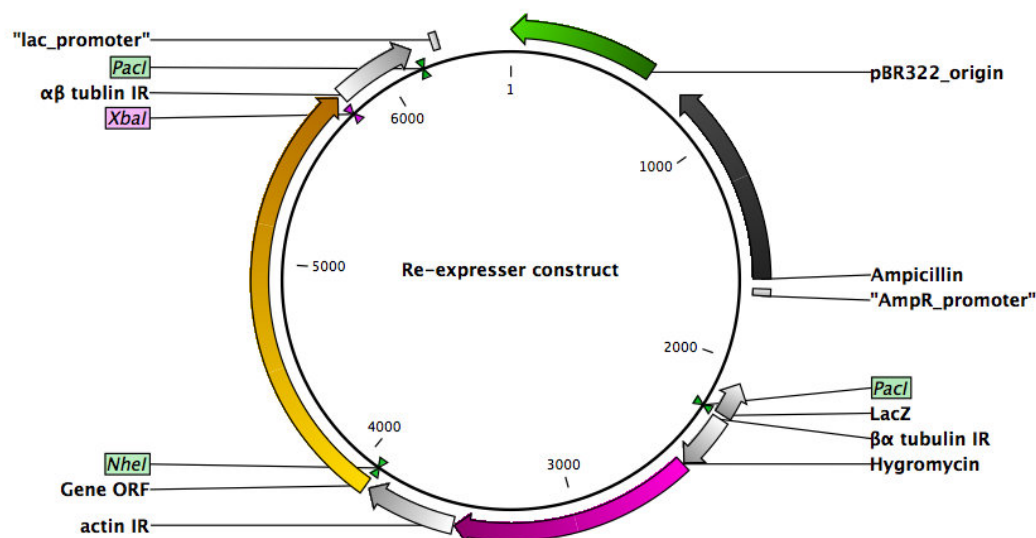
### 3.5.1 Generation of re-expresser constructs

After confirming the viability and assaying the DNA damage phenotype of the *recq2*, *mus81* and *pif6* mutants, it was necessary to re-express the genes in the mutant backgrounds to confirm that the phenotype was due to the knockout of the gene and not as a result of secondary effects. To achieve this the entire ORF of each gene was amplified (primers *RECQ2* #151/#126; *MUS81* #152/#128; *PIF6* #153/#130, see Appendix 7.1 for sequences) by PCR and cloned, using *NheI* and *XbaI* restriction sites, into a modified version of the pGL2070 vector (Gift, Jeremy Mottram). pGL2070 contains a blasticidin resistance cassette, flanked upstream with a  $\beta\alpha$ -tubulin intergenic region and downstream with an actin intergenic region ( $\beta\alpha$ -BSD-act). The re-expressed ORF is placed downstream of the actin intergenic region and is followed immediately downstream by a 6HA tag and an  $\alpha\beta$ -tubulin intergenic region ( $\beta\alpha$ -BSD-act-ORF-6HA- $\alpha\beta$ ). pGL2070 was modified by replacing the blasticidin resistance gene with hygromycin (using cloning primers #108/#109, see Appendix 7.1), followed by removal of the 6HA tag and replacement with an oligonucleotide (oligos #116/#117, see Appendix 7.1) containing multiple restriction sites, including *NheI* and *XbaI*, which were used to attempt to clone in the *RECQ2*, *MUS81* and *PIF6* ORFs. Due to the compatibility of *NheI* and *XbaI* digested DNA ends, plasmids were screened by restriction enzyme digestion prior to sequencing in order to exclude clones in which the ORF had ligated in the reverse orientation.

A representation of the expected final plasmids used to generate the re-expresser cell lines is shown in Figure 3-20. The plasmid contains a hygromycin resistance cassette for selection and the untagged gene ORF. This construct can be integrated into the  $\beta\alpha$  tubulin locus of *T. brucei* after digestion with *PacI* and transformation. Recombination occurs on the  $\beta\alpha$  and  $\alpha\beta$  targeting regions, integrating the construct into the  $\beta\alpha$ -tubulin array replacing an  $\alpha$  tubulin ORF and allowing expression of the ORF mRNA through the flanking non-endogenous intergenic sequences.

It was not possible to generate a *RECQ2* re-expresser plasmid; the successfully amplified ORF did not successfully ligate into the construct. The *MUS81* and *PIF6*

re-expresser constructs RXP-MUS81 (re-expresser MUS81) and RXP-PIF6 (re-expresser PIF6) were digested with *PacI* and transformed into either the *pif6*<sup>-/-</sup> or *mus81*<sup>-/-</sup> cell lines and selected for using 2.5 µg.mL<sup>-1</sup> G418, 5 µg.mL<sup>-1</sup> blasticidin and 10 µg.mL<sup>-1</sup> hygromycin.



**Figure 3-20 Schematic of re-expresser construct**

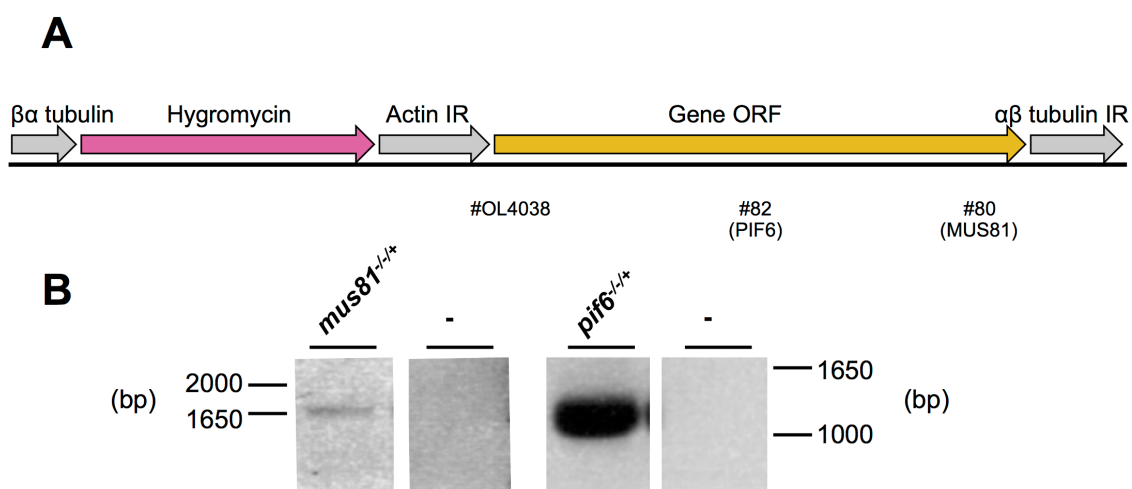
The MUS81 and PIF6 re-expresser constructs used to generate the re-expresser lines in *mus81*<sup>-/-</sup> and *pif6*<sup>-/-</sup> backgrounds contain the MUS81 (1845 bp) or PIF6 ORF (2391 bp) (orange), cloned using *NheI* and *XbaI* restriction enzyme sites (RECQ2 primers #151 and #126; MUS81 primers #152 and #128; PIF6 primers #153 and #130). Antibiotic selection is enabled by a hygromycin resistance gene (pink). The construct is linearised by digestion with *PacI* (sites indicated) and integration into the tubulin array for expression of the resistance gene and re-expressed ORF is facilitated by tubulin sequences flanking the construct. Tubulin and actin sequences present provide transplicing and polyadenylation signals for hygromycin and the re-expressed ORF. Sizes shown (bp) vary by gene ORF.

Two of the three *mus81*<sup>-/-/+</sup> transformation tissue cultures plates contained too many positive wells to be clonal. However, the third and most dilute plate contained six wells with live cells. These were all taken for analysis and PCR amplification showed all six clones contained the MUS81-RXP cassette, as discussed below. Two of the three *pif6*<sup>-/-/+</sup> transformation tissue cultures plates contained too many positive wells to be clonal. However, the third and most dilute plate contained nine wells with live cells. Six were taken for analysis and PCR amplification showed all six clones contained the PIF6-RXP cassette, as discussed below.

### 3.5.2 Confirmation of re-expresser cell lines by PCR and RT-PCR

Possible *MUS81* and *PIF6* re-expresser clones following transformation of *mus81*<sup>-/-</sup> and *pif6*<sup>-/-</sup> cell lines with the RXP-MUS81 or RXP-PIF6 re-expresser constructs

respectively ( $\beta\alpha$ -HYG-Act-ORF- $\alpha\beta$ ) were screened using PCR primers that bind to the actin IR (primer #OL4038, see Appendix 7.1) and within the gene ORF (primer #80, *MUS81* (1603 bp product); and #82, *PIF6* (1165 bp PCR product); see Appendix 7.1). Figure 3-21A shows the PCR screening strategy. The re-expresser construct was amplified in the *MUS81* and *PIF6* re-expresser cell lines (Fig. 3-21B), demonstrating the successful integration of the re-expresser constructs into the *mus81*<sup>-/-</sup> and *pif6*<sup>-/-</sup> cell lines. The putative re-expresser cells were designated *mus81*<sup>-/-/+</sup> and *pif6*<sup>-/-/+</sup>.

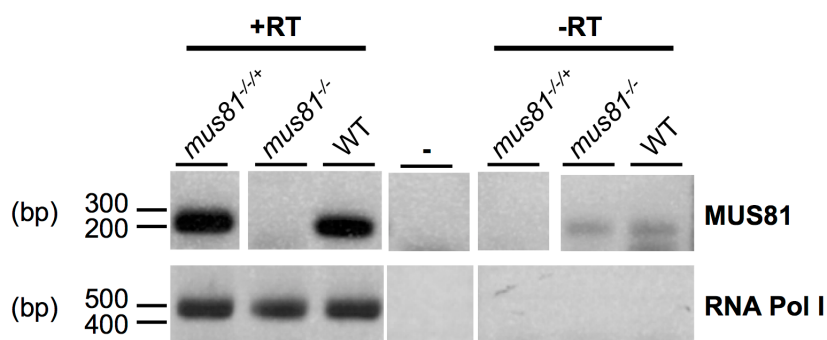


**Figure 3-21** PCR confirmation of *MUS81* and *PIF6* re-expresser cell lines (A) Integration of the RXP-*MUS81* and RXP-*PIF6* constructs was confirmed by PCR amplification of the construct using a forward primer binding in the acting intergenic region (primer #OL4038) and reverse primer binding in the *MUS81* (primer #80, 1603 bp product) of *PIF6* (primer #82, 1165 bp product) ORF. (B) PCR amplification of the re-expresser construct was performed as described above, using genomic DNA extracted from *mus81*<sup>-/-/+</sup> and *pif6*<sup>-/-/+</sup> cells. Distilled water was used as a negative control. Sizes are shown (bp).

PCR analysis of cDNA from putative *mus81*<sup>-/-/+</sup> and wild type cells was undertaken to test for re-expression of *MUS81*. A ~200 bp region of the *MUS81* ORF was PCR amplified (primers #79/#80, see Appendix 7.1). A negative reverse transcriptase (-RT) control reaction, which contained no enzyme, was performed to test whether PCR products generated arose from cDNA and not from contaminating genomic DNA. In wild type, *mus81*<sup>+/-</sup> and *mus81*<sup>-/-/+</sup> both *MUS81* and RNA PolI PCR products were amplified (Fig. 3-22) confirming that cDNA is present in all samples and that *MUS81* is expressed in all four cell lines. Absence of PCR amplification of either PCR product in the minus RT samples of *mus81*<sup>-/-/+</sup> confirms the absence of detectable genomic DNA contamination in the RNA sample used for cDNA synthesis. However, amplification in the -RT sample of wild type cells indicates some genomic DNA contamination in the RNA sample



used for cDNA synthesis. RT-PCR analysis was not undertaken for the putative *pif6*<sup>-/-/+</sup> line.

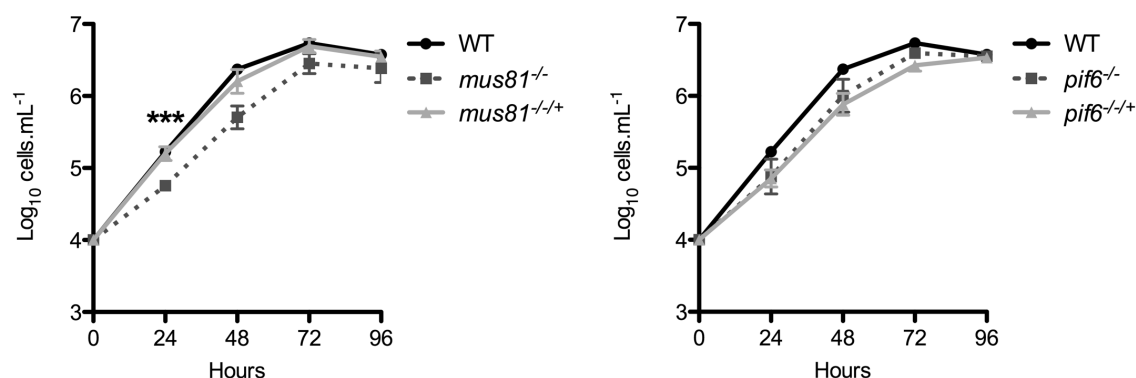


**Figure 3-22 RT-PCR confirmation of MUS81 re-expresser cell lines**  
cDNA was synthesised from RNA extracted from wild type, *mus81*<sup>-/-</sup> and *mus81*<sup>-/-/+</sup> cells (+RT). A 214 bp region of the *MUS81* ORF was amplified by PCR using primers #79 and #80, and a 453 bp region of RNA polymerase I (RNA pol I) was PCR amplified to test for the presence of cDNA in all samples. PCR amplification was also carried out on samples from a reaction in which reverse transcriptase (-RT) was absent. Distilled water was used as a negative control (-). Size and markers are shown (ladder, bp).

## 3.6 Analysis of re-expresser cell lines

### 3.6.1 Analysis of *in vitro* growth of re-expresser cell lines

Analysis of *in vitro* growth of re-expresser lines was performed as described in Section 3.4.1. The growth of *mus81*<sup>-/-/+</sup> cells showed growth similar to that of wild type (Fig. 3-23A), indicating a rescue of the growth defect of *mus81*<sup>-/-</sup> cells. *In vitro* growth of *pif6*<sup>-/-/+</sup> cells was indistinguishable from *pif6*<sup>-/-</sup> cells (Fig. 3-23B) with both showing some evidence of slowed, though not statistically significant, growth relative to wild type. Whether this means that there is no re-expression of PIF6 is unclear since previous growth curves, as well as MMS and HU survival assays, suggest that *pif6*<sup>+/-</sup> cells are phenotypically equivalent to *pif6*<sup>-/-</sup> cells. For this reason, re-introduction of a single copy of the *PIF6* ORF may have little discernible effect.



**Figure 3-23** *In vitro* growth of *MUS81* and *PIF6* re-expresser cell lines

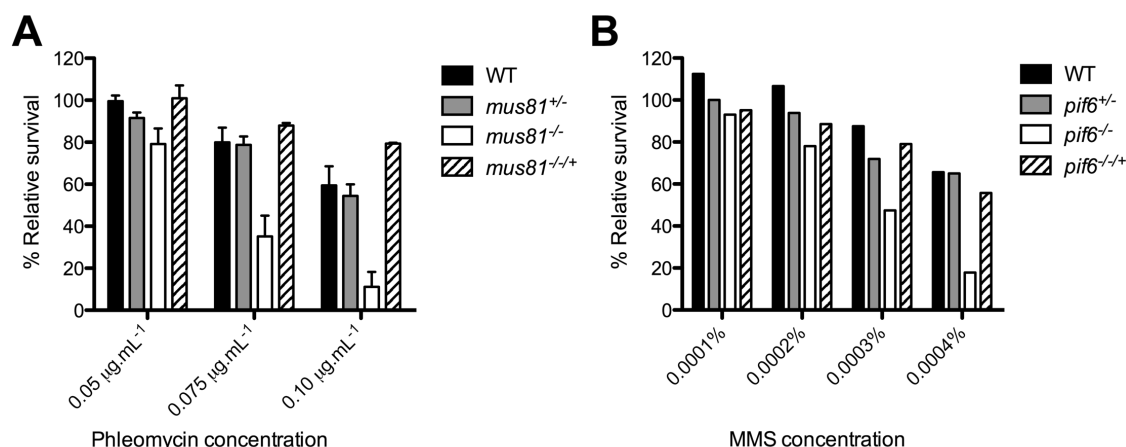
Wild type, *mus81*<sup>-/-</sup> and *mus81*<sup>-/-/+</sup> cells were incubated for 96 hours at a starting density of 1 x 10<sup>4</sup> cells.mL<sup>-1</sup>. The cell density of cultures was counted every 24 hours using a haemocytometer. Cultures were not diluted during incubation. Growth curves were repeated three times and the mean cell density is plotted logarithmically. Bars represent standard error of the mean (SEM). Differences between the mean cell density of WT and *mus81*<sup>-/-</sup> cultures were calculated using Student's t-test; \*\*\* p=0.005).

### 3.6.2 Analysis of DNA damage sensitivity of re-expresser cell lines

To test further for functionality of the re-expressed *MUS81*, rescue of the DNA damage phenotype using a clonal survival in the presence of phleomycin was investigated next. This assay was chosen due to the strong increase in phleomycin sensitivity in *mus81*<sup>-/-</sup> mutants. Re-expression of the *MUS81* ORF in the *mus81*<sup>-/-/+</sup> line clearly rescued the phleomycin sensitivity of *mus81*<sup>-/-</sup> mutants (Fig. 3-24A).

Clonal survival in the presence of MMS was used to attempt to ascertain whether re-introduction of the *PIF6* ORF showed phenotypic evidence. However, as stated above, this was complicated by the fact that the observed increased resistance of *pif6*<sup>-/-</sup> mutants to MMS relative to wild type was also seen in *pif6*<sup>+/-</sup> cells (Fig. 3-17). Even accepting these problems however, the results of the assay (Fig. 3-24B) were anomalous when compared with previous clonal survival assays of *pif6* mutants in the presence of MMS; in this experiment *pif6*<sup>-/-</sup> mutants displayed substantially impaired survival compared with wild type and *pif6*<sup>+/-</sup> cells, casting doubt on the validity of the experiment. The survival of *pif6*<sup>-/-/+</sup> cells following MMS treatment was similar to wild type, though as discussed above, whether this suggests that the expressed *PIF6* ORF in *pif6*<sup>-/-/+</sup> rescues the phenotype of *pif6*<sup>-/-</sup> mutants is unclear.

The Pro742Ser mutation present in the re-expressed protein, combined with the absence of RT-PCR data to demonstrate re-expression of *PIF6*, mean it is not possible to determine whether *PIF6* re-expression alters the DNA damage phenotype, if it has been rescued, or indeed whether it is expressed at all. As the *pif6* mutants showed statistically significant increased sensitivity to HU, it is possible that a clonal survival assay with this compound may be revealing, but again the *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> mutants appear phenotypically indistinguishable.



**Figure 3-24 Clonal survival of re-expressor lines in the presence of DNA damaging agents**

(A) Wild type (WT), *mus81*<sup>+/-</sup>, *mus81*<sup>-/-</sup> and *mus81*<sup>-/-+</sup> cells were diluted to the equivalent of one cell per well over three 96 well plates in the presence of 0/0.05/0.075/0.1 µg.mL<sup>-1</sup> phleomycin. Plates were incubated for 10 days or until wells were clearly positive or negative by colour (saturated cultures are indicated by yellow media). Positive wells were counted and the survival of phleomycin-treated cultures was calculated relative to untreated. Mean survival (%) per plate relative to untreated cells is plotted; bars represent standard error of the mean (SEM) from the three plates, to indicate intra-experiment variation. (B) Wild type, *pif6*<sup>+/-</sup>, *pif6*<sup>-/-</sup> and *pif6*<sup>-/-+</sup> cells were treated and analysed as in (A), with the exception that cells were treated with 0%, 0.0001%, 0.0002%, 0.0003% or 0.0004% MMS.

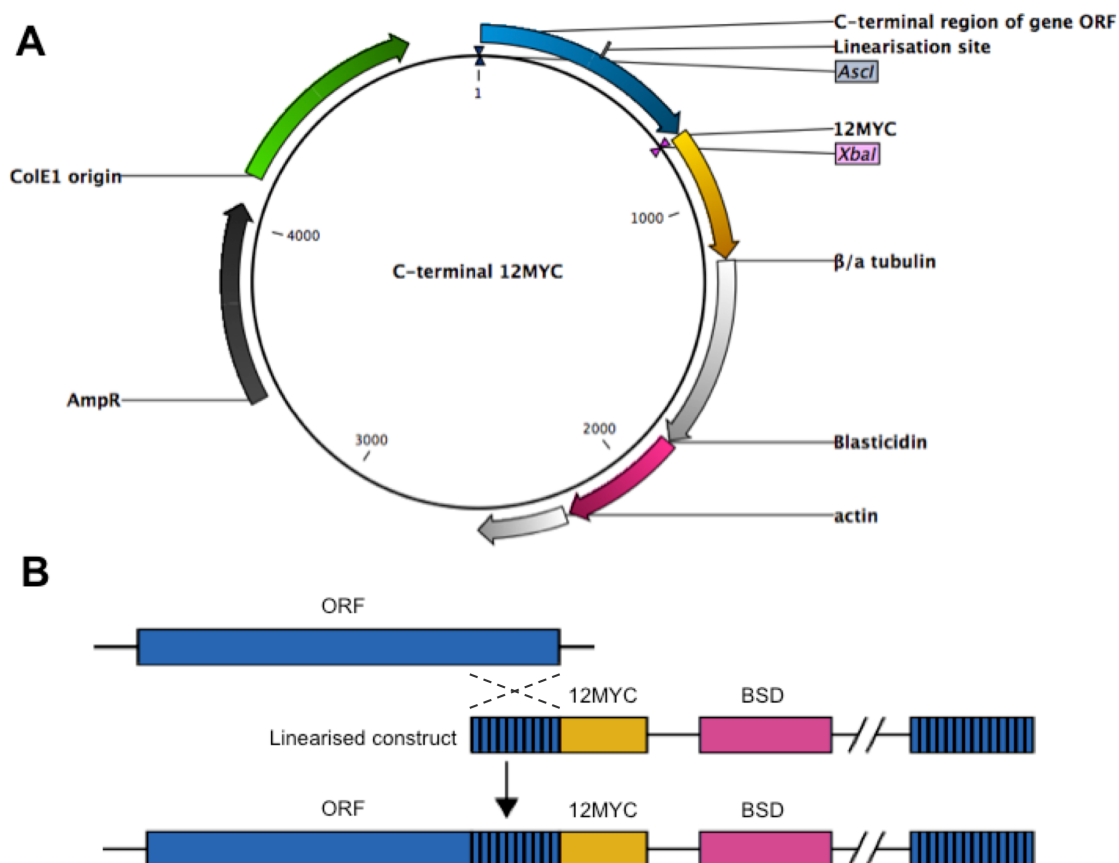
### 3.7 Generation of endogenously tagged cell lines

#### 3.7.1 Generation of C-terminal myc-tagged MUS81 and PIF6 cell lines

To date there are no antisera available against any of these three *T. brucei* repair proteins and so, in order to attempt to discern their subcellular localisation, attempts were made to express each of them as epitope tagged variants. It was initially intended to endogenously tag all three factors at the C-terminus using 12 copies of the myc epitope. This vector (pNAT<sup>x12myc</sup>, (Alsford & Horn, 2008)) has successfully been used to tag and localise proteins in *T. brucei* previously (Jones *et al.*, 2014; Trenaman *et al.*, 2013). However, cloning to

generate a C-terminal *RECQ2::12myc* construct was unsuccessful and this protein was instead tagged at the N-terminus (see below).

C-terminal tagging strategy and the C-terminal tagging constructs are shown in Figure 3-25. The *MUS81* ORF was targeted using the final 488 bp of the C-terminal coding region of the gene containing a unique restriction site (EcoRV) in the *MUS81::C-12myc* construct (Gift, Catarina Marques), derived from the pNAT<sup>x12myc</sup> construct (Alsford & Horn, 2008). The *PIF6* ORF was targeted using the final 743 bp of the C-terminal coding region of the gene containing a unique restriction site (BsiWI) in the *PIF6::C-12myc* construct, derived from the pNAT<sup>x12myc</sup> construct (Alsford & Horn, 2008). The *PIF6* targeting region was PCR amplified from wild type genomic DNA using primers #19 and #20 (see Appendix 7.1). The unique restriction enzyme site was used to linearise the construct prior to transformation. This C-terminal region includes sequence up to but not including the stop codon to allow translation to proceed through into the 12myc epitope tag, and was cloned into the construct using HindIII and XbaI restriction enzyme sites. Selection of transformants is enabled by a blasticidin resistance gene, which is flanked by tubulin and actin processing sequences. After integration the 3' UTR of the tagged gene will no longer be endogenous sequence; instead, plasmid sequence is downstream and the blasticidin cassette provides polyadenylation signals for the tagged gene.



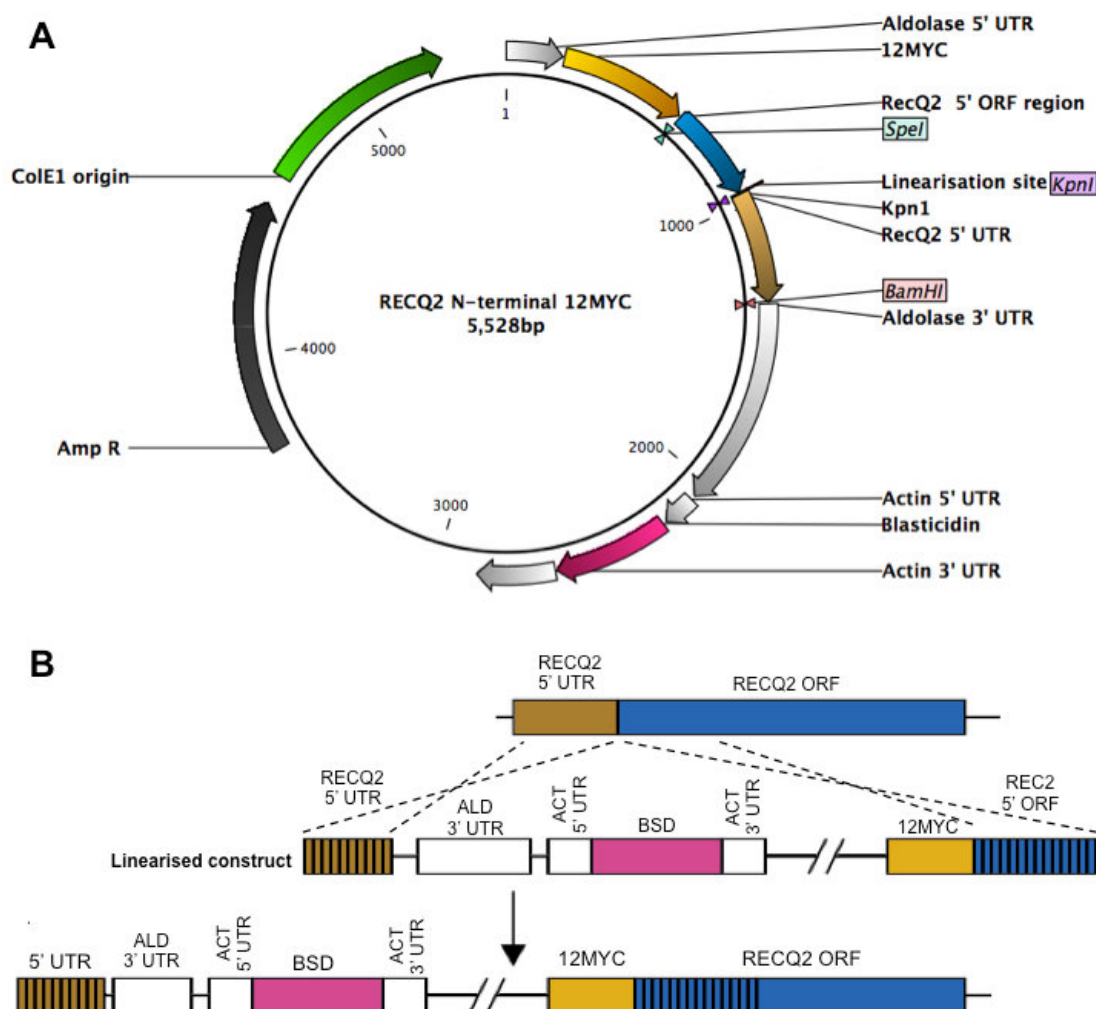
**Figure 3-25** Diagram of C-terminal 12myc tagging construct and strategy  
 (A) Construct used to epitope tag MUS81 and PIF6 at the C-terminus with a 12myc tag. As described in the text, the final 488 bp of MUS81 for the *MUS81::C-12myc* construct and final 743 bp of PIF6 for the *PIF6::C-12myc* construct was cloned, excluding the stop codon (blue), immediately upstream of 12 repeats of the myc sequence (orange) using *Ascl* and *Xba* restriction enzyme sites. Antibiotic selection was provided by a blasticidin resistance gene (pink), flanked by tubulin and actin sequences for mRNA transplicing and polyadenylation (light grey). The *MUS81::C-12myc* and *PIF6::C-12myc* constructs were linearised using a unique restriction enzyme linearisation site located in the cloned C-terminal region of the gene ORF (indicated on map), *EcoRV* for *MUS81::C-12myc* and *BsiWI* for *PIF6::C-12myc*.  
 (B) C-terminal tagging strategy used. Linearisation of the *MUS81::C-12myc* or *BsiWI* for *PIF6::C-12myc* construct within the cloned 3' end of the ORF (blue dashed box) allows integration into the genome at the site of the targeted ORF (blue) by HR (dashed lines). This produces the endogenous ORF fused with the C-terminal 12myc tag (orange). Not to scale.

### 3.7.2 Generation of N-terminal myc-tagged RECQ2 cell line

The strategy used to generate the RECQ2 endogenous N-terminal 12myc-tagged lines and the constructs used are shown in Figure 3-26. A 322 bp region of the *RECQ2* ORF, starting after the start codon, was cloned into the construct immediately downstream and in frame of the 12myc tag. The forward ORF primer (#24) contained a *SpeI* restriction enzyme site for ligation of the PCR product to the construct backbone, and the reverse primer (#25) contained a *KpnI* restriction enzyme site for ligation of the *RECQ2* ORF PCR product to the *RECQ2* UTR PCR product (see below). The cloned *RECQ2* ORF region is

immediately upstream of a 378 bp region of the *RECQ2* 5' UTR, which includes the *RECQ2* 5' UTR region up to but excluding the *RECQ2* start codon. The forward 5' *RECQ2* UTR primer (#22) contained a KpnI restriction enzyme site for ligation of the *RECQ2* UTR PCR product to the *RECQ2* ORF PCR product and the reverse primer (#23) contained a BamHI restriction enzyme site for ligation of the *RECQ2* UTR PCR product to the construct backbone. The sequences of all primers can be found in Appendix 7.1.

The 5' UTR and ORF regions facilitated targeting of *RECQ2* through homologous recombination and integration into the genome. The KpnI restriction site at the junction between the ORF and 5' UTR was used for linearisation prior to transformation. Selection of transformants is facilitated by a blasticidin resistance gene, flanked 5' by an aldolase processing sequence and 3' by an actin processing sequence. After integration of the construct, the 5' UTR of the *12myc::RECQ2* gene is non-endogenous due to the use of the sequences of the upstream blasticidin resistance gene as RNA trans-splicing sequences.



**Figure 3-26** RECQ2 N-terminal 12myc tagging construct and strategy

(A) Construct used to epitope tag RECQ2 at the N-terminus with a 12myc tag. As described in the text, the first 322 bp of the *RECQ2* ORF, excluding the stop codon (blue), was cloned immediately downstream of 12 repeats of the myc sequence (orange) using *SpeI* and *KpnI* restriction enzyme sites. Immediately downstream of the cloned *RECQ2* ORF region is a 378 bp region of the 5' UTR (brown) of *RECQ2* up to but not including the *RECQ2* start codon, using restriction enzyme sites *KpnI* and *BamHI*. The [12myc-*RECQ2* ORF- 5' UTR] cassette is flanked by aldolase 5' and 3' UTRs (grey) for mRNA transplicing and polyadenylation. Antibiotic selection is enabled by a blasticidin resistance gene, flanked by actin sequences for transplicing and polyadenylation. Restriction enzyme digestion using *KpnI* was used to linearise the construct prior to transformation. Sizes shown (bp).

(B) N-terminal tagging strategy used. Linearisation (using *KpnI*) of the *RECQ2-Nmyc* construct allows integration into the genome at the site of the targeted ORF (blue) by HR (dashed lines). This produces the endogenous ORF fused with the N-terminal 12myc tag (orange). BSD, blasticidin; ALD, aldolase. Not to scale.

### 3.7.3 Confirmation of *in vivo* myc-tagging by western blot analysis

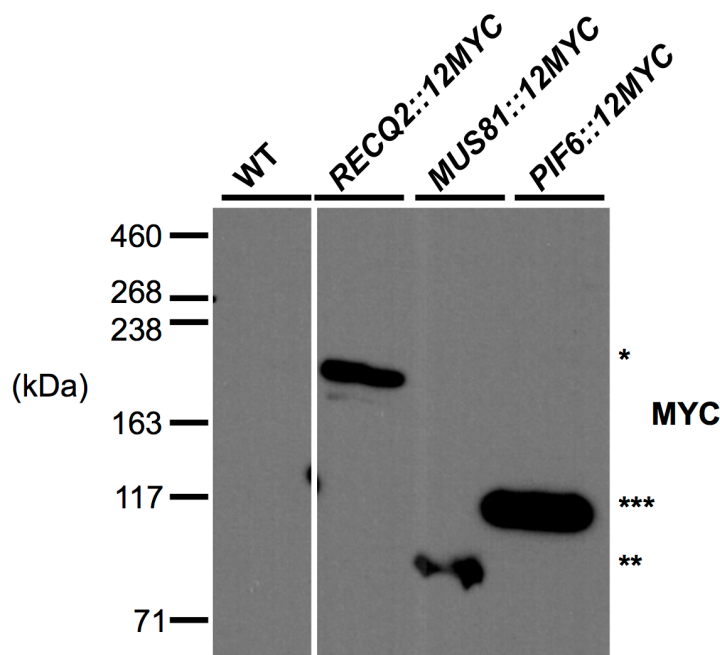
Wild type bloodstream Lister 427 cells were transformed with the linearised *MUS81::12myc*, *PIF6::12myc* and *12myc::RECQ2* constructs in order to generate a C-terminally 12myc tagged *MUS81* and *PIF6*, and N-terminally 12myc tagged *RECQ2* cell lines. Antibiotic resistant transformants were selected using

10  $\mu\text{g} \cdot \text{mL}^{-1}$  blasticidin. Six putative RECQ2 12myc clones were analysed by dot blot for myc expression, three of which expressed the myc epitope (data not shown). Whole cell lysates were dotted onto nitrocellulose membrane and probed with rabbit anti-myc antiserum (1:7000 dilution, Life Technologies) and secondary anti-rabbit antiserum (1:5000, Life Technologies). The clone selected for use in experiments was later analysed by western blot (Fig. 3-27) to test that the myc-tagged protein expressed matched the expected molecular weight of RECQ2 12myc. Six putative PIF6 12myc clones were analysed by western blot, of which four expressed a myc-reactive protein matching the expected molecular weight of PIF6 12myc (data not shown). Eighteen putative MUS81 12myc clones were analysed by PCR due to problems with dot and western blots at the time of clone screening. This PCR used a forward primer (#52) binding within the *MUS81* ORF upstream of the unique restriction enzyme site and a reverse primer (#64) binding downstream of the 12myc tag, PCR amplifying a 1196 bp region to test integration of the *MUS81* 12myc construct. Four of the eighteen putative *MUS81* 12myc clones were positive for this PCR (data not shown). The clone selected for use in experiments was analysed by western blot (Fig. 3-27) to test that *MUS81* 12myc was expressed.

Putative myc-tagged RECQ2, *MUS81* and PIF6 transformant clones were screened by western blot to test expression of the tagged proteins. Whole cell lysates were separated by SDS-PAGE on a 3-8% tris-acetate gel before western blotting and probing with rabbit anti-myc antiserum (1:7000 dilution, Millipore) and secondary anti-rabbit HRP-conjugated antiserum (1:5000 dilution, Life Technologies). Figure 3-27 demonstrates that the putative myc-tagged clones expressed the expected tagged protein. The predicted molecular weight of 12myc RECQ2 is 182 kDa (167.8 kDa (RECQ2) + 14.4 kDa (12myc)), and a single band was observed at approximately this molecular weight. The predicted molecular weight of *MUS81* 12myc is 82.6 kDa (68.2 kDa (*MUS81*) + 14.4 kDa (12myc)) and a single band was observed at approximately this molecular weight. The predicted molecular weight of PIF6 12myc is 102.7 kDa (88.3 kDa (PIF6) + 14.4 kDa (12myc)) and a single band was observed with an apparent molecular weight of approximately 117 kDa, larger than predicted. Using a 10% bis-tris gel, the apparent molecular weight of PIF6 12myc was closer to the expected molecular weight, suggesting that the smaller size here is a running



artefact for PIF6 in 3-8% tris-acetate gels. These myc-tagged clones were used in all subsequent experiments involving tagged RECQ2, MUS81 or PIF6.



**Figure 3-27 Western blot of 12myc-tagged RECQ2, MUS81 and PIF6**  
Whole cell lysates ( $5 \times 10^6$  cell equivalents) of untagged wild type (WT), RECQ2 12 myc, MUS81 12myc and PIF6 12myc cultures were separated by SDS-PAGE on a 3-8% tris-acetate gel, western blotted and probed with rabbit anti-Myc antiserum (Millipore, 1:7000 dilution) and secondary anti-mouse HRP-conjugated antiserum (Life Technologies, 1:5000 dilution). The bands produced by RECQ2 12myc (182 kDa), MUS81 12myc (82.6 kDa) and PIF6 12myc (102.7 kDa) are indicated (\*, \*\* and \*\*\* respectively) and size markers are shown (kDa).

### 3.8 Testing the functionality of endogenously-tagged proteins

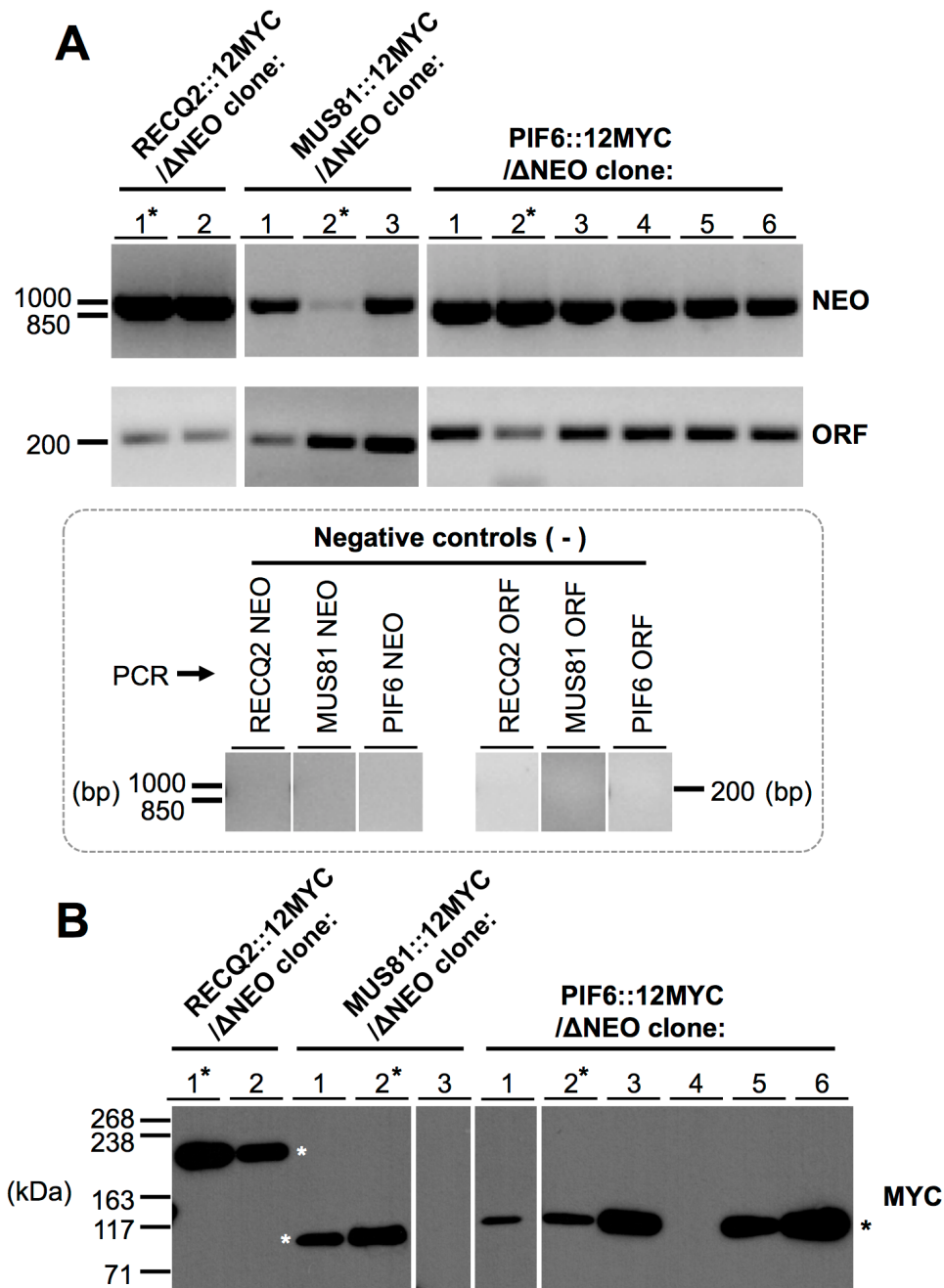
#### 3.8.1 Generation and validation of single-copy myc tagged cell lines

As it had already been established that *recq2*, *mus81* and *pif6* null mutants are viable, the functionality of the tagged proteins in the 12myc lines could not be determined merely by showing that it is possible to knockout out the second, untagged allele of the genes. Instead, for each tagged line, the second gene allele was knocked out and the DNA damage phenotype of the resulting cell lines was assayed.

Because the myc tagging in all cases was selected for by blasticidin resistance, the second (untagged) allele was knocked out utilizing the three different neomycin knockout constructs used to generate the knockout cell lines (see

Section 3.3). Each knockout construct was transformed into the cognate 12myc cell lines and transformants were selected with  $5 \mu\text{g.mL}^{-1}$  blasticidin and  $2.5 \mu\text{g.mL}^{-1}$  G418. Putative 12myc/- clones were screened using PCR and western blotting. A PCR to confirm integration of the neomycin knockout cassette into the correct locus ("NEO PCR") and a PCR reaction amplifying a ~200 bp region of the targeted ORF (as a positive control; "ORF PCR") (see Section 3.3.3) were both performed. Figure 3-28 shows that all of the putative 12myc/- clones analysed for each of RECQ2, MUS81 and PIF6 were positive for both of these PCRs, indicating correct integration of the neomycin knockout cassette in the clones.

Two clones of each cell line, previously shown to be correct by PCR (as above), were screened by western blotting to confirm they still expressed the tagged protein. Whole cell lysates were separated by SDS-PAGE on a 3-8% tris-acetate gel before western blotting and probing with anti-myc antiserum at a 1:7000 dilution (Millipore). As shown in Figure 3-28, all of the clones tested expressed the 12myc tagged protein for that line. One clone from each cell line (indicated by \* ) was selected for use in further experiments.



**Figure 3-28 Confirmation of 12myc<sup>-/-</sup> cells by western blot and PCR**  
**(A)** Genomic DNA extracted from putative 12myc RECQ2<sup>-/-</sup>, MUS81 12myc<sup>-/-</sup> and PIF6 12myc<sup>-/-</sup> clones and PCR amplification was used to confirm integration of the  $::\Delta$ NEO knockout construct (NEO) and presence of the targeting gene ORF, as used previously to check integration of knockout cassettes. NEO PCR: RECQ2, primers #51 and #155 1008 bp; MUS81, primers #52 and #155 974 bp; PIF6, primers #53 and #155 898 bp. ORF PCR: RECQ2, primers #77 and #78 232 bp; MUS81, primers #79 and #80 214 bp; PIF6, primers #81 and #82 245 bp. Primer sequences can be found in Appendix 7.1. Distilled water was used for negative controls for all PCRs (gels in dashed box). Ladder and sizes indicated (bp). Dashed lines between gels indicate different gels aligned for this figure. **(B)** Whole cell lysates ( $5 \times 10^6$  cell equivalents) of 12myc RECQ2<sup>-/-</sup>, MUS81 12myc<sup>-/-</sup> and PIF6 12myc<sup>-/-</sup> cultures were separated by SDS-PAGE on a 3-8% tris-acetate gel, western blotted and probed with mouse anti-myc antiserum (Millipore, 1:7000 dilution) and secondary anti-mouse HRP-conjugated antisera (Life Technologies, 1:5000 dilution). The bands produced by RECQ2 12myc (182 kDa), MUS81 12myc (82.6 kDa) and PIF6 12myc (102.7 kDa) are indicated (\*, \*\* and \*\*\* respectively) and size markers are shown (kDa).

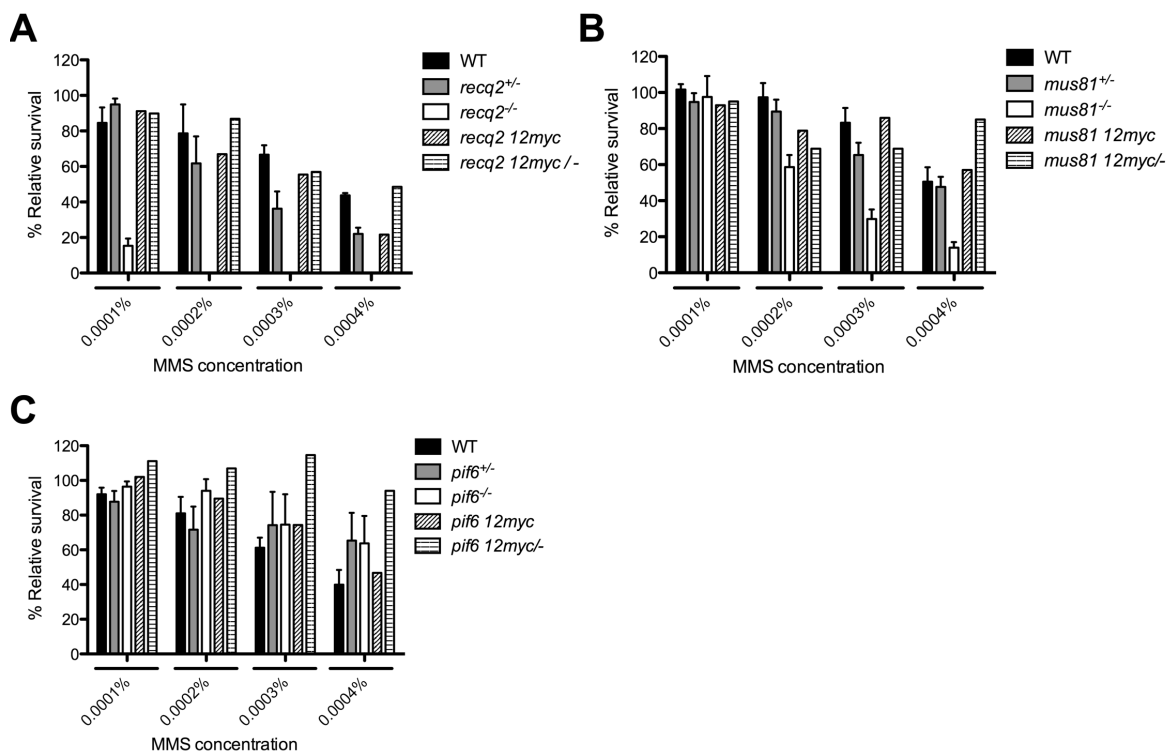
### 3.8.2 Testing functionality using DNA damage sensitivity assays

The functionality of the myc-tagged proteins was evaluated using a clonal survival assay in the presence of MMS. MMS was the damaging agent of choice as previous data (Section 3.4.3.1) had indicated *recq2*<sup>-/-</sup>, *mus81*<sup>-/-</sup> and *pif6*<sup>-/-</sup> cells all exhibit an altered clonal survival phenotype in its presence; decreased survival in *recq2*<sup>-/-</sup> and *mus81*<sup>-/-</sup>, versus increased survival in *pif6*<sup>-/-</sup>.

The clonal survival of the *recq2 12myc*<sup>-/-</sup> cells was similar to wild type cells at all concentrations of MMS (Fig. 3-29A). If the myc-tagged RECQ2 were non-functional, the cell line would be expected to have the same strong sensitivity to MMS as *recq2*<sup>-/-</sup>. As this is clearly not the case, these data demonstrate that 12myc RECQ2 is functional in MMS repair. It was apparent from previous assays that *recq2*<sup>+/-</sup> cells also display some increased MMS sensitivity relative to wild type cells. Whether the greater comparability of the *recq2 myc*<sup>-/-</sup> cells' survival to that of wild type vs *recq2*<sup>+/-</sup> indicates some increased protein expression due to the tagging is not possible to test in the absence of antiserum against native RECQ2. In addition, since the same effect was not seen in the cells in which one RECQ2 allele was tagged and the other unaltered (12myc RECQ2), it is possible that these data merely reflect the inherit variability in this assay. Similarly, *mus81 12myc*<sup>-/-</sup> cells did not display the same increased sensitivity to MMS as *mus81*<sup>-/-</sup> mutants (Fig. 3-29B), indicating that MUS81 12myc is also functional in MMS repair. Interpretation of the data for PIF6, however, is more complex.

As discussed earlier (Section 3.4.3.2), both *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> cells consistently exhibit a mild increase in MMS sensitivity at high concentrations (especially 0.0004%) of MMS compared to wild type, though this difference is not statistically significant. When assayed, *pif6 12myc* (with one tagged allele, and one functional untagged allele) cells exhibited survival similar to wild type at 0.0004% MMS, while *pif6 12myc*<sup>-/-</sup> survival was much higher than wild type and also higher than *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> (Fig. 3-29C). Alone, increase in MMS resistance of *pif6 12myc*<sup>-/-</sup> suggests that PIF6 12myc is non-functional as the phenotype is most similar to *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> mutants. However, if this were the case, *pif6 12myc* would also be expected to exhibit the same increase in MMS resistance as *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> are both more resistant to MMS. Since the data do not conform to this predicted result, it is not possible to conclude

whether PIF6 12myc is functional. Using HU as the damaging agent in this experiment may be more appropriate given HU treatment results in a distinct phenotype in *pif6*<sup>-/-</sup> mutants.



**Figure 3-29 Clonal survival of myc/- cells following MMS treatment**  
Wild type (WT) and +/-, -/-, 12myc and 12myc/- *recq2* (A), *mus81* (B) and *pif6* (C) mutants were diluted to the equivalent of one cell per well over three 96 well plates in the presence of varying concentrations of phleomycin. Plates were incubated for 10 days or until wells were clearly positive or negative by colour (saturated cultures are indicated by yellow media). Positive wells were counted and the survival of MMS-treated cultures was calculated relative to untreated. Mean survival (%) relative to untreated cells is plotted; bars represent standard error of the mean (SEM).

### 3.9 Localisation of endogenously-tagged factors

*T. brucei* cells generated expressing 12myc RECQ2, MUS81 12myc or PIF6 12myc at the endogenous locus (Sections 3.7-3.8) with the second allele untagged were used for microscopic protein localisation by indirect immunofluorescence. DNA was stained with DAPI and visualised using a DAPI filter; 12myc-tagged proteins were immunolocalised using Alexa Fluor 488 conjugated anti-myc antiserum (Life Technologies, 1:2000 dilution) and visualised using a FITC filter. Where co-localisation with RAD51 was assessed, RAD51 was immunolocalised using an anti-Rad51 antiserum (1:1000 dilution) and Alexa Fluor 594 conjugated anti-rabbit antiserum (Life Technologies, 1:2000 dilution), and visualised using a rhodamine

filter. Whole cells were visualised using DIC microscopy. See Materials and Methods, Section 2.10 for full details.

### 3.9.1 Localisation of 12myc-RECQ2

In the vast majority of cells, no signal could be detected for 12myc RECQ2 (Fig.3-30F). However, in a very small proportion (0.2%) of cells, discrete, punctate signals were seen in the nucleus (images not shown). This pattern is reminiscent of the localisation described for *T. brucei* RAD51 (Dobson *et al.*, 2011; Glover *et al.*, 2008; Hartley & McCulloch, 2008; Proudfoot & McCulloch, 2005; Trenaman *et al.*, 2013), which is normally not seen in the cell, but localises in what have been described as foci in a small number of cells in the absence of induced damage. RAD51 foci are thought to be repair-related structures as their number increases after damage, both in *T. brucei* and in many other cells (Bergink *et al.*, 2013; Da Ines *et al.*, 2013; Haaf *et al.*, 1995; Tarsounas *et al.*, 2004). Therefore, 12myc RECQ2 cells were also probed for RAD51 using anti-RAD51 antiserum (see Materials and Methods, Section 2.10 for full details) to examine RAD51/12myc RECQ2 co-localisation. RAD51 foci were observed in ~2% of untreated cells (Fig. 3-30F), similar to previous studies (Dobson *et al.*, 2011; Hartley & McCulloch, 2008; Proudfoot & McCulloch, 2005; Trenaman *et al.*, 2013) and slightly higher than 12myc RECQ2 (~1% of untreated cells), though this may simply reflect the very small number of cells that display either localisation. As a control, untagged wild type cells were probed with anti-myc antiserum. No signal was observed in these control cells. Untagged wild type cells were also probed with anti-rabbit Alexa Fluor 594 antiserum to test for cross-reactivity. No signal was observed in these cells.

To assess whether the 12mycRECQ2 signals might represent repair-related foci and, indeed, might be structurally associated with RAD51 foci, the cells were incubated for 18 hours with phleomycin at  $1\text{ }\mu\text{g.mL}^{-1}$ , which has been shown to generate a majority of cells with RAD51 foci (Dobson *et al.*, 2011; Hartley & McCulloch, 2008; Trenaman *et al.*, 2013). In these conditions greatly increased numbers of both 12myc RECQ2 and RAD51 foci were observed; 47% of cells contained one or more 12myc RECQ2 foci and 44% of cells contained one or more RAD51 foci (377 cells analysed). The relative proportion of cells with 12myc RECQ2 and RAD51 foci were very similar, as was the pattern of foci

accumulation: most cells contained a single focus of either protein, though some contained two or three discrete foci, while others had larger numbers (where it became difficult to count accurately). Beyond this striking similarity in the behaviour of each factor, there was evidence for overlapping signals, though a number of different localisation patterns were seen, examples of which are shown in Figures 3-30B-E: cells containing a single RECQ2 12myc focus that co-localised with a single RAD51 focus (Fig. 3-30B), a 12myc RECQ2 focus with no accompanying RAD51 focus (Fig. 3-30C); multiple 12myc RECQ2 and RAD51 foci that partially or fully co-localised (Fig. 3-30D); and one or more RAD51 foci with no 12myc RECQ2 foci (Fig. 3-30E). Overall, ~50% of cells contained one or more 12myc RECQ2 foci with most of these containing only one focus (Fig. 3-30G). The percentage of cells containing RAD51 foci and the number of foci per cell were similar to that of 12myc RECQ2 (Fig. 3-30H). The majority of 12myc RECQ2/RAD51 foci co-localised fully (~60% of cells with both foci, Fig. 3-30H), though a smaller number localised only partially (~20% of cells with foci, Fig. 3-30h). In summary, there appears to be substantial, but not complete overlap in behaviour between 12myc RECQ2 and RAD51, especially after phleomycin-induced damage. Whether the non-overlapping signals merely reflect incomplete resolution of one or other signal or tell us that the proteins act in subtly different manners (perhaps temporally or spatially) is unclear.

**Figure 3-30 Representative examples of 12myc RECQ2 localisation**  
 12myc RECQ2 cells were incubated for 18 hours in the presence or absence of phleomycin ( $1 \mu\text{g.mL}^{-1}$ ). Cells were fixed and Alexa Fluor 488 antiserum was used to stain RECQ2 12myc; RAD51 was localised using anti-RAD51 antiserum and goat anti-rabbit Alexa Fluor 594 conjugated antiserum (Life Technologies). DAPI was used to visualise the DNA (DAPI) as described in the text. Differential interference contrast was used to visualise whole cells (DIC). DAPI, 12myc RECQ2, RAD51 and merged (DAPI, 12myc RECQ2 and RAD51) images are shown. (A-E) Representative images of different types of 12myc RECQ2 and RAD51 localisation, as described in the text. (F, G) The percentage of cells containing 12myc RECQ2 and RAD51 foci and the number of foci, in the absence (G) and presence (H) of phleomycin, was counted using Fiji ImageJ. (H) Cells containing 12myc RECQ2 and RAD51 foci following phleomycin treatment were categorised according to the degree of 12myc RECQ2 and RAD51 foci co-localisation, represented as % of cells that contained both foci. Bar, 13  $\mu\text{M}$ .

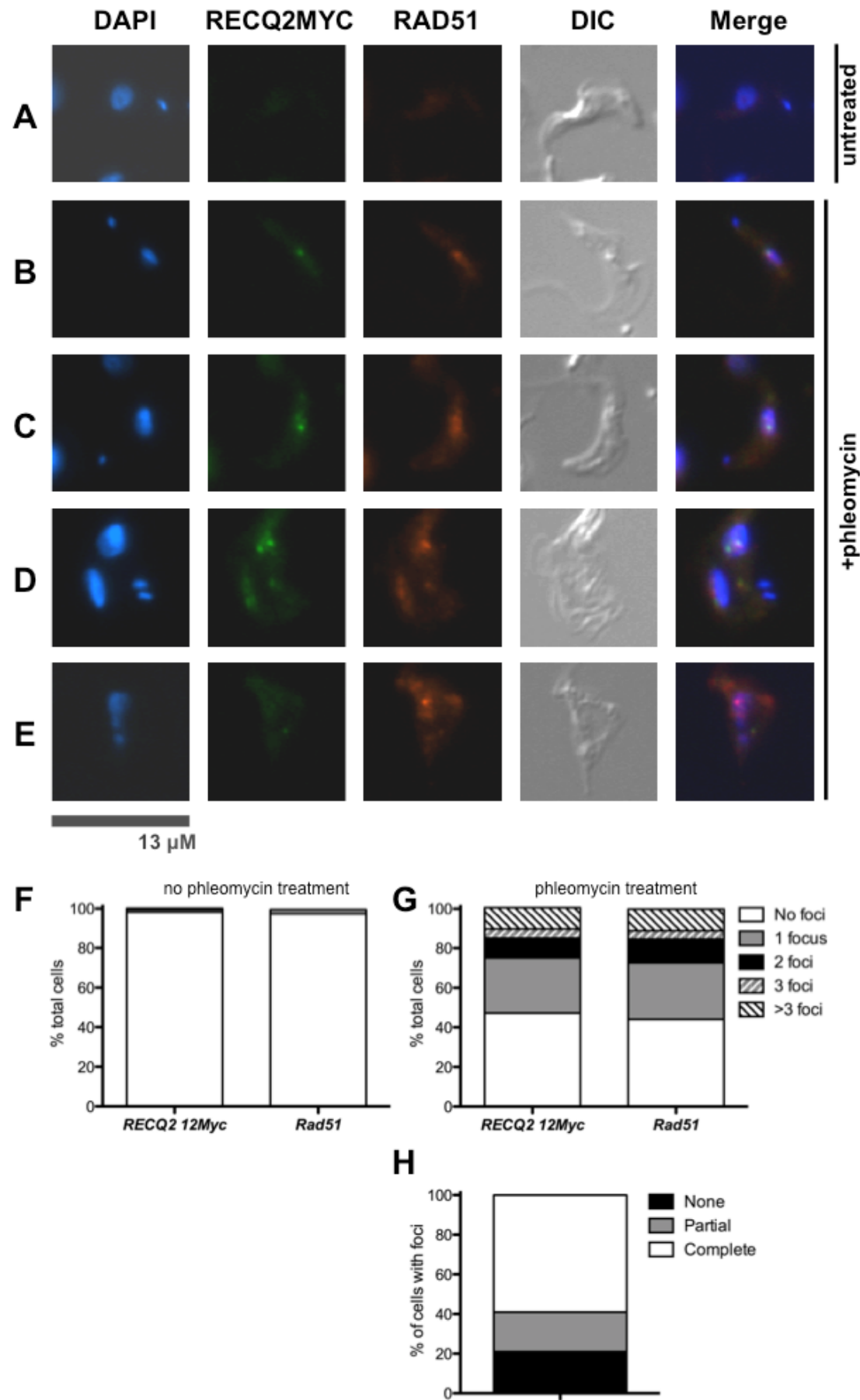
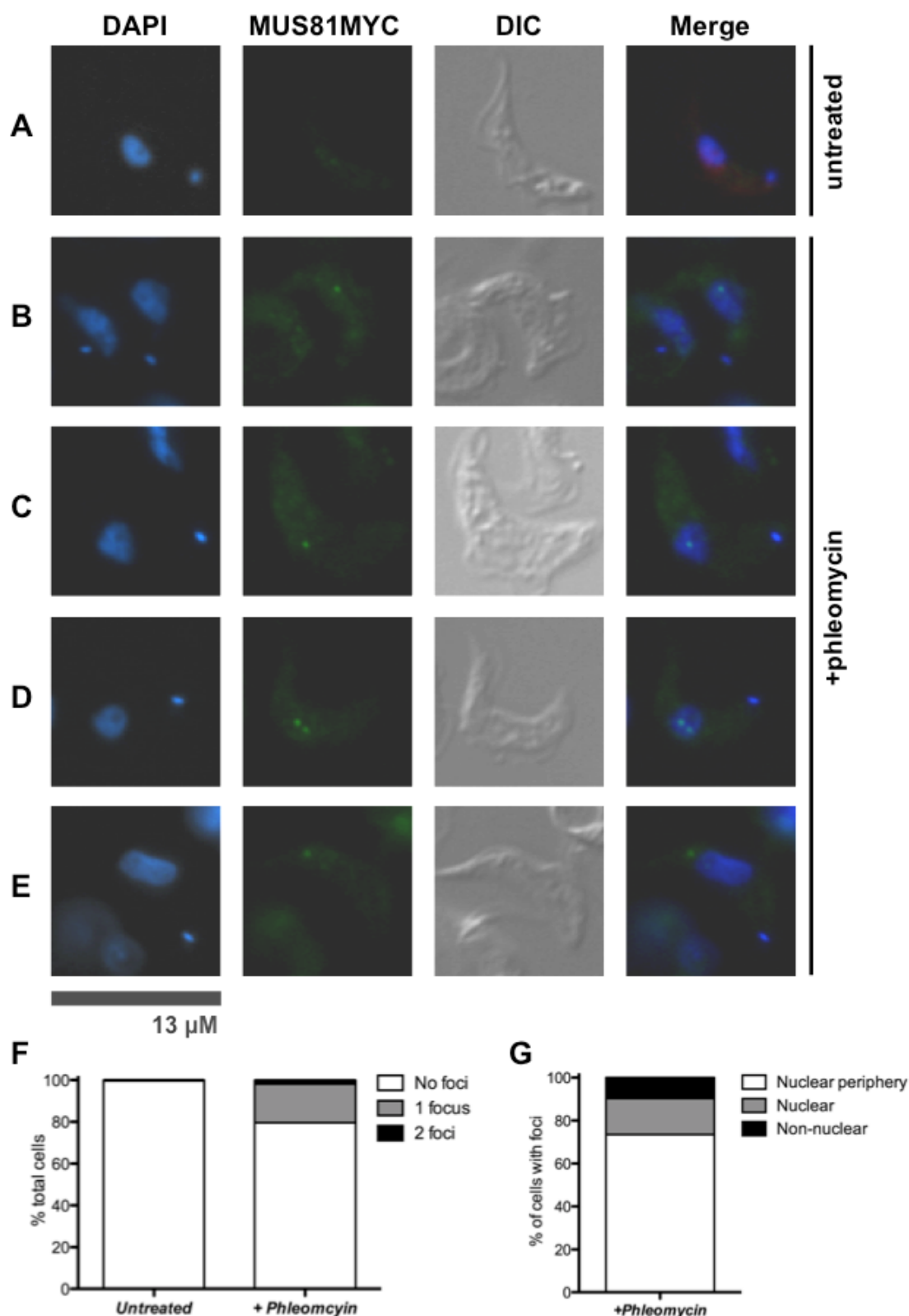


Figure 3-30



### 3.9.2 Localisation of MUS81myc

Similarly to 12myc RECQ2, <1% of cells contained detectable MUS81 nuclear signal in the absence of DNA damage (Fig. 3-31A) with the huge majority of cells showing no fluorescence. Where signal was seen, it was again in discrete foci-like structures. Therefore, the cells were incubated with phleomycin using the same conditions. In this case, however, only approximately 20% of cells contained MUS81 12myc foci (Fig. 3-31F), and in almost all cases the treated cells with MUS81 12myc foci contained just a single focus. As this pattern of focal accumulation differs from 12mycRECQ2 and RAD51, no attempt at co-localisation of MUS81 12myc with RAD51 was performed. Instead, the distribution of localisation of the foci was analysed, revealing four different types of localisation, examples of which are shown in Figure 3-31B-E. In the majority (73%) of cells (Fig. 3-31G), MUS81 12myc foci were located in the nuclear periphery (Fig. 3-31B). A further 17% of cells (Fig. 3-31G) contained foci more centrally located in the nucleus, as shown in Figure 3-31C. Fewer than 2% of cells (18 out of 998 cells counted) contained two nuclear foci, as shown in Figure 3-31D. MUS81 12myc foci (Fig. 3-31E) were also infrequently observed (approximately 10% of cells with foci, Fig. 3-31G) that did not overlap with the DAPI. Though these might be non-nuclear, it is also possible that these are in nuclear locations in which the chromatin has disassembled, perhaps due to the DNA damage.



**Figure 3-31** Representative examples of MUS81 12myc localisation  
 MUS81 12myc cells were incubated for 18 hours in the presence or absence of phleomycin (+T/-T) ( $1 \mu\text{g} \cdot \text{mL}^{-1}$ ). Cells were fixed and anti-myc Alexa Fluor 488 (Millipore) antiserum was used to detect MUS81 12myc. DAPI was used to visualise the DNA (DAPI) as described in the text. Differential interference contrast was used to visualise whole cells (DIC). DAPI, MUS81 12myc and merged DAPI and MUS81 12myc images are shown. (A-E) Representative images of different types of MUS81 12myc localisation as described in the text. (F) Percentage of cells containing MUS81 12myc foci. The number of foci present, in the absence and presence of phleomycin, was counted using Fiji ImageJ. (G) Cells

containing MUS81 12myc foci following phleomycin treatment were categorised by sub-cellular localisation, represented as % of cells that contained MUS81 12myc foci. Bar, 13  $\mu$ M.

### 3.9.3 Localisation of PIF6myc

In contrast to 12myc RECQ2 and MUS81 12myc, PIF6 12myc signal was observed in the nucleus of approximately 65% of cells (Fig. 3-32E) in the absence of DNA damage. An example of a PIF6 12myc positive cell is shown in Figure 3-32A, while Figure 3-32B shows a PIF6 12myc negative cell. PIF6 12myc nuclear localisation is in agreement with the findings of Liu *et al.* (2009a), who reported nuclear localisation of GFP-tagged PIF6, albeit in procyclic form cells. What distinguishes those BSF cells that display a nuclear PIF6 12myc signal from those that do is unclear. Again, in contrast with 12myc RECQ2 and MUS81 12myc, foci of PIF6 12myc were not observed after incubation with the DNA damaging agent phleomycin using the same conditions. However, the level of PIF6 12myc fluorescence appeared to be stronger after phleomycin (Fig. 3-32C) and the number of PIF6 12myc positive cells observed increased by 8% (Fig. 3-32E). Rarely, phleomycin-treated cells were observed with diffuse, grainy PIF6 12myc staining (Fig. 3-32D). It is possible this is due to a high background level of immunofluorescence or the breakdown of the nucleus in dying cells.

Given the perceived increase in nuclear PIF6 12myc fluorescence of phleomycin treated cells, nuclear fluorescence intensity was measured using ImageJ. The mean fluorescence intensity was measured for each nucleus and the mean of all the cells then calculated. The mean nuclear fluorescence intensity of untagged cells was used as a measure of the background level of fluorescence (autofluorescence). The nuclear fluorescence intensity of PIF6 12myc cells was higher than that of untagged cells (Fig. 3-32F), suggesting that the signal is derived from PIF6 12myc. Surprisingly, the mean fluorescence intensity of both PIF6 12myc nuclei and untagged nuclei was higher following phleomycin treatment (Fig. 3-32F). It is unclear why phleomycin treatment would increase nuclear autofluorescence, though metabolic changes induced by phleomycin treatment could affect autofluorescence if they altered the levels of autofluorescent molecules such as NADH and NADPH (Monici, 2005). However, after subtraction of background fluorescence (untagged nuclear fluorescence) the nuclear fluorescence intensity of phleomycin treated PIF6 12myc cells is

higher than that of untreated PIF6 12myc cells (Fig. 3-32G), appearing to confirm the damage-dependent change in PIF6 12myc signal. The basis for such a change is unknown, but could reflect an increase in PIF6 expression in response to phleomycin damage, which might be consistent with a role for the factor in repair of such damage. If so, why *pif6* mutants do not display altered sensitivity to phleomycin relative to wild type cells is unclear.

Since PIF6 12myc expression appeared not to be consistent in all nuclei, it was asked whether the expression of PIF6 12myc varied by the cell cycle stage. Cell cycle dependent expression could suggest that the PIF6 is only involved in the repair of DNA at certain cell cycle stages. The average nuclear fluorescence of untreated nuclei was measured as described above; nuclei were categorised by the cell cycle stage of the cell (1N1K, 1N2K or 2N2K) and plotted as a dot plot (Fig. 3-32H). Error bars represent standard error of the mean (SEM). The intensity of PIF6 12myc expression appeared to have no obvious correlation with the cell cycle stage, with the mean fluorescence 6.8 in 1N1K, 7.5 in 1N2K and 5.6 in 2N2K. The mean PIF6 12myc fluorescence was lowest in 2N2K cells and the spread of values also smaller, with no values exceeding 10, but this could be the result of the smaller number of 2N2K cells in the population compared to 1N1K and 1N2K. It should also be borne in mind when considering these data that, unlike for RECQ2 and MUS81, it remains unclear if the myc tagging of PIF6 affects the function (and thus perhaps localisation) of the factor.

**Figure 3-32 Representative examples of PIF6 12myc localisation**

PIF6 12myc cells were incubated for 18 hours in the presence or absence of phleomycin ( $1 \mu\text{g} \cdot \text{mL}^{-1}$ ). Fixed cells were fixed and Alexa Fluor 488 (Millipore) was used to stain PIF6 12myc. DAPI was used to visualise the DNA (DAPI) as described in the text. Differential interference contrast was used to visualise whole cells (DIC). DAPI, PIF6 12myc and merged DAPI and PIF6 12myc images are shown. (A-D) Representative images of different types of PIF6 12myc localisation, as described in the text. Bar,  $13 \mu\text{M}$ . (E) The percentage of cells containing nuclear PIF6 12myc nuclear signal, in the absence and presence of phleomycin, was counted using Fiji ImageJ. (F) The average nuclear myc fluorescence intensity was measured in wild type untagged and PIF6 12myc cell lines in the absence of and following phleomycin treatment. Fluorescence intensity was measured using Fiji ImageJ, arbitrary units. (G) The nuclear myc fluorescence intensity of PIF6 12myc cells, with and without phleomycin treatment, above the background level. The background level of fluorescence (fluorescence of untagged cell nuclei (F)) was subtracted from the tagged nuclei fluorescence (fluorescence of PIF6 12myc, (F)). (H) Average nuclear PIF6 12myc fluorescence by cell cycle stage. The average fluorescence intensity of nuclei was measured using ImageJ and plotted according to cell cycle stage (1N1K, 1N2K, 1N2K). Horizontal lines on H represent mean, and bars the standard error of the mean (SEM).

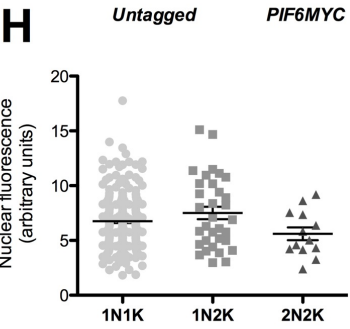
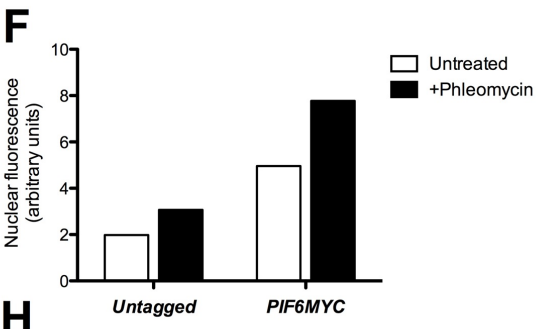
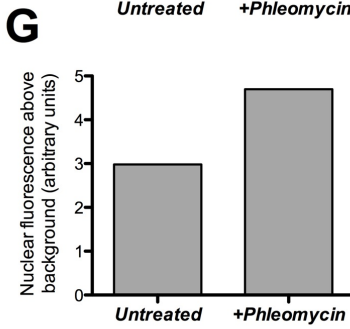
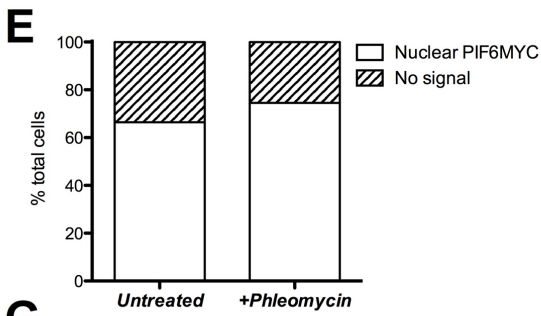
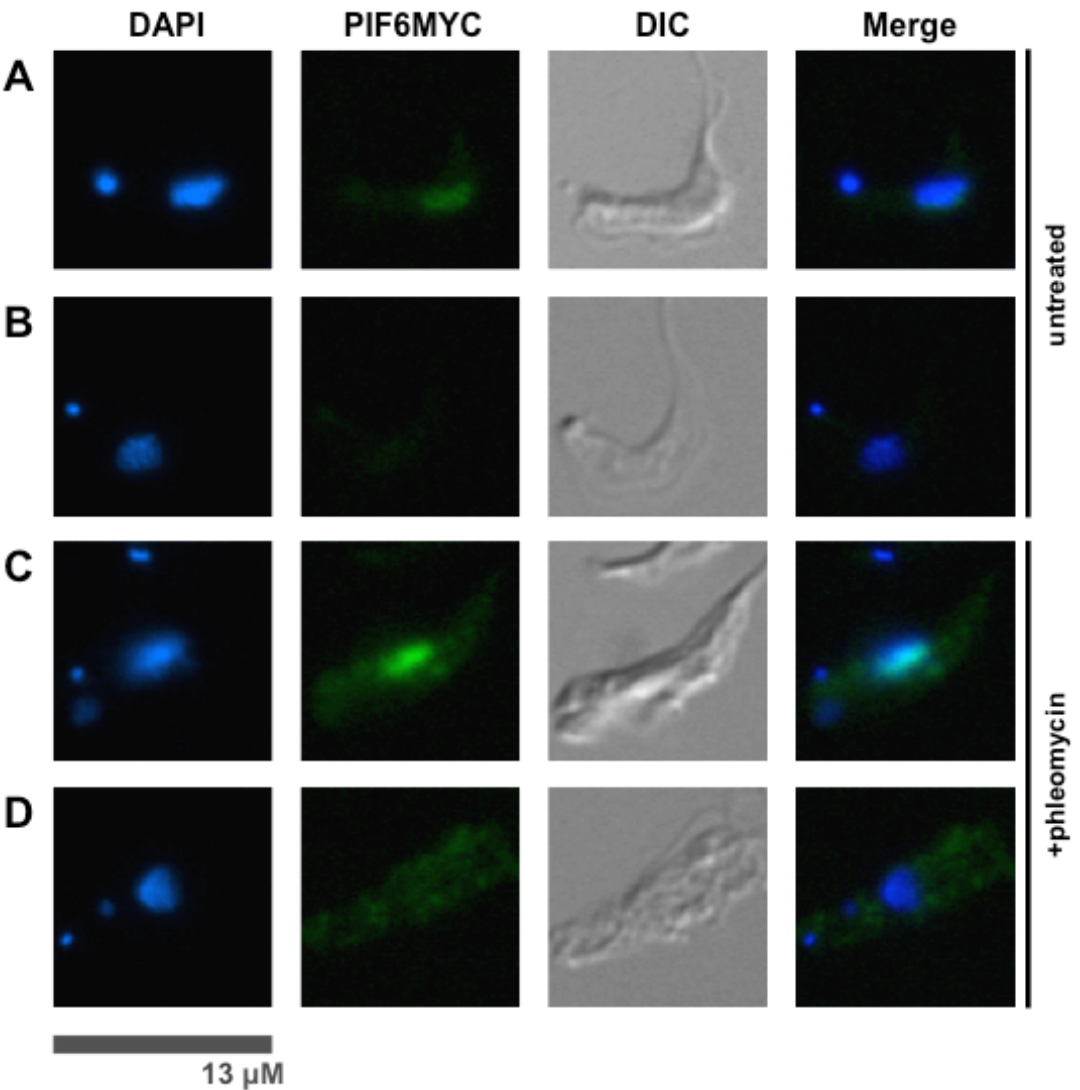


Figure 3-32

## 3.10 Summary

### 3.10.1 Key findings of this chapter

The data in this chapter has revealed at least two new factors that act in *T. brucei* genome repair: RECQ2 and MUS81. The data presented also suggest that a third factor, PIF6 may also be involved in the repair of DNA damage in *T. brucei*. Analysis of gene mutants, as well as localisation of the myc-tagged factors by gene mutation and localisation, suggests differences in their function.

#### 3.10.1.1 *T. brucei* RECQ2

Analysis of *recq2* mutants revealed a range of defects in growth and sensitivity to DNA damaging agents. *recq2*<sup>-/-</sup> mutants displayed impaired growth compared to wild type cells, as did *recq2*<sup>+/-</sup> to some extent, though the effect was much less pronounced. In assays measuring sensitivity to DNA damaging agents a similar pattern was observed, with *recq2*<sup>+/-</sup> mutants behaving similarly to wild type. *In vitro* growth analysis revealed pronounced sensitivity to MMS and some evidence for sensitivity to phleomycin in *recq2*<sup>-/-</sup> cells compared to wild type, both of which were also observed when clonal survival assays were used to assess DNA damaging agent sensitivity. Each of these agents cause DNA breaks, predominantly DSBs in the case of phleomycin (Groth *et al.*, 2010; Reiter *et al.*, 1972) and via nucleotide alkylation in the case of MMS (Groth *et al.*, 2010; Tercero & Diffley, 2001). These data suggest that RECQ2 has repair roles in response to these damaging agents.

The increased sensitivity to MMS in *recq2*<sup>-/-</sup> mutants is consistent with the MMS treatment phenotype of other RecQ-like helicase mutants. MMS sensitivity has previously been reported in other RecQ helicase mutants. *S. cerevisiae sgs1* mutants (Mullen *et al.*, 2000) as well as human *blm*<sup>-/-</sup> (Wang *et al.*, 2000) and chicken DT40 cells (Imamura *et al.*, 2001) are hypersensitive to MMS. Human *recql1*<sup>-/-</sup> and *recql5*<sup>-/-</sup> cells however, do not show any difference in MMS sensitivity compared to wild type cells (Wang *et al.*, 2003). There was no information found in the current literature on the phleomycin sensitivity of other RecQ-like helicase mutants. Therefore, comparison of these data with phleomycin sensitivity of other RecQ helicase mutants cannot be made.

A mild growth defect was observed in *recq2*<sup>-/-</sup> mutants following HU treatment and a strong sensitivity was seen in *recq2*<sup>-/-</sup> mutants in clonal survival assays. HU depletes the cellular dNTP pool, resulting in replication stalling (Bianchi *et al.*, 1986), which can lead to DNA breaks if the replication is not successfully restarted. This suggests that RECQ2 may have a role at the replication fork. Other RecQ helicases have been demonstrated to function at the replication fork: BLM and WRN contribute to normal replication fork progression, and treatment with HU in BLM and WRN mutants results in defects in replication fork recovery (Mao *et al.*, 2010; Sidorova *et al.*, 2013). Additionally, BLM and WRN mutants accumulate a higher level of DNA damage than wild type cells following HU treatment, as measured by  $\gamma$ H2A staining (Mao *et al.*, 2010).

Localisation of RECQ2 using a 12myc RECQ2 cell line and immunofluorescent microscopic analysis was unsuccessful. This could be due to low expression of the protein. However, following treatment with phleomycin, 12myc RECQ2 foci were observed. This is reminiscent of RAD51, which is not normally seen but, following DNA damage, localises to discrete nuclear foci - thought to be repair-related structures (Dobson *et al.*, 2011; Hartley & McCulloch, 2008; Proudfoot & McCulloch, 2005; Trenaman *et al.*, 2013). Co-staining of 12myc RECQ2 and RAD51 following phleomycin damage revealed substantial but not complete overlap in RAD51 and 12myc RECQ2 following phleomycin damage, suggesting that 12myc RECQ2 foci are repair-related structures.

RECQ2 has been proposed as the *T. brucei* homologue of BLM and the third component of the three-protein RTR (RecQ-Top3-Rmi1) complex (Kim & Cross, 2011). The complex is composed of a RecQ helicase (BLM in mammals, SGS1 in yeast), a topoisomerase III  $\alpha$  and RMI1 (RecQ mediated genome instability 1; also known as BLAP75/18) (Johnson *et al.*, 2000; Wu *et al.*, 2000; Yin *et al.*, 2005). The RTR complex is a major regulator of mitotic crossover in eukaryotes and resolves double HJs in a non-crossover manner (Bussen *et al.*, 2007; Chelysheva *et al.*, 2008; Raynard *et al.*, 2008). Kim and Cross identified the *T. brucei* RMI1 and TOPO3 $\alpha$  homologues, demonstrated that they interact and characterised null mutants (Kim & Cross, 2010; Kim & Cross, 2011), but the RecQ helicase component of the complex is yet to be identified and characterised. *T. brucei* *rmi1*<sup>-/-</sup> and *topo3a*<sup>-/-</sup> mutants both display a mild growth impairment similar to that of *recq2*<sup>-/-</sup> mutants described here. Also in common with *recq2*<sup>-/-</sup> mutants,

*rmi1*<sup>-/-</sup> and *topo3a*<sup>-/-</sup> mutants display increased HU sensitivity. Furthermore, *topo3a*<sup>-/-</sup> mutants display markedly increased phleomycin sensitivity, which was likewise observed in *recq2*<sup>-/-</sup> mutants albeit not with as strong sensitivity.

As discussed above (Section 3.2.1), the *T. brucei* genome encodes two RecQ-like helicases: *TbRECQ2* (Tb427.08.6690) and *TbRECQ1* (Tb427.06.3580), the latter of which is discussed in Chapter 5. Protein domain prediction, multiple sequence alignment and phylogenetic analysis showed that RECQ2 has greater similarity to BLM and SGS1 homologues than RECQ1. Combined with the similarity of the DNA damage and growth phenotypes of *recq2*<sup>-/-</sup> mutants to *rmi1*<sup>-/-</sup> and *topo3a*<sup>-/-</sup> mutants, the data presented in this chapter suggest that RECQ2 may be the RecQ helicase component of the RTR complex in *T. brucei*. *rmi1*<sup>-/-</sup> and *topo3a*<sup>-/-</sup> mutants also display an elevated VSG switching rate and increased crossover switching events (Kim & Cross, 2010; Kim & Cross, 2011). It was therefore decided to investigate the VSG switching profile of *recq2* mutants, described in the next chapter (Chapter 4).

### 3.10.1.2 *T. brucei* MUS81

Generation and analysis of *mus81*<sup>+/-</sup> and *mus81*<sup>-/-</sup> mutants revealed that similar to *recq2*<sup>-/-</sup> mutants, *mus81*<sup>-/-</sup> mutants have a growth defect. An increase in phleomycin and MMS sensitivity in *mus81*<sup>-/-</sup> mutants in clonal survival assays suggests that *T. brucei* MUS81 is involved in the response to both phleomycin and MMS damage. Re-expression of the *MUS81* ORF in *mus81*<sup>-/-</sup> cells from the tubulin locus (*mus81*<sup>-/-/+</sup> cells) restored the phleomycin sensitivity to wild type levels. In contrast to RECQ2 however, clonal survival data suggests that MUS81 plays little or no role in the response to HU.

The DNA damage phenotype of *mus81* mutants described in this chapter is consistent with reports of yeast MUS81. Increased sensitivity of *mus81*<sup>-/-</sup> cells to MMS and phleomycin is consistent with observations in *S. cerevisiae mus81Δ* mutants (Blanco *et al.*, 2010; Interthal & Heyer, 2000). Furthermore, though following HU treatment in *S. cerevisiae*, MUS81 is required for recovery from replication fork collapse (which was not examined here), mutants do not have altered viability after HU treatment (Kai *et al.*, 2005), nor altered survival in human cells (Nomura *et al.*, 2007).



MUS81 12myc immunofluorescence signal was not observed in BSF *T. brucei* cells in the absence of phleomycin treatment. However, following treatment with phleomycin, MUS81 12myc foci were observed (in ~90% of cases) in the nucleus. This is reminiscent of both the appearance of nuclear RAD51 (Dobson *et al.*, 2011; Hartley & McCulloch, 2008; Proudfoot & McCulloch, 2005; Trenaman *et al.*, 2013) and 12myc RECQ2 foci (Section 3.9.1) following DNA damage and suggests that these MUS81 12myc foci are associated with DNA repair. Additionally, data indicated that MUS81 12myc is functional, as MUS81 12myc/- cells displayed MMS sensitivity similar to wild type rather than *mus81*<sup>-/-</sup> cells.

Similar to BSF *T. brucei* MUS81, human MUS81 is predominantly nuclear, with super accumulation in nucleoli and co-localising with BLM and WRN RecQ-like helicases in nucleoli (Gao *et al.*, 2003). Yeast MUS81 also localises to the nucleus (Fu & Xiao, 2003). Though no literature was available on the localisation of MUS81 in other eukaryotes following phleomycin treatment, human MUS81 is recruited to regions of UV damage in human S-phase cells, but not cells that have completed replication (Gao *et al.*, 2003).

In summary, the data presented in this chapter suggest that although primary amino acid sequence analysis indicated that *T. brucei* MUS81 appears to be somewhat diverged from other eukaryotic MUS81 homologues, it shares similar DNA damage sensitivity and localisation phenotypes with yeast and human MUS81 endonucleases. MUS81 appears to be involved in the response to phleomycin and MMS, though not HU, and further, MUS81 12myc forms what may be DNA repair-related foci in response to phleomycin damage.

### 3.10.1.3 *T. brucei* PIF6

Analysis of *pif6* mutants in this chapter revealed that unlike *recq2*<sup>-/-</sup> and *mus81*<sup>-/-</sup> mutants, neither *pif6*<sup>+/-</sup> nor *pif6*<sup>-/-</sup> mutants display a marked *in vitro* growth defect. The DNA damage phenotype of *pif6* mutants was markedly different to *recq2* and *mus81* mutants. Though no defect was observed in the growth of *pif6* mutants following MMS treatment, both *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> mutants displayed an increase in MMS resistance in clonal survival assays. The opposite effect was observed following HU treatment; *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> mutants displayed a mild growth defect and an increase in HU sensitivity in clonal survival assays. PIF6

does not appear to be involved in the response to phleomycin, displaying only a minor growth defect in *pif6*<sup>+/-</sup> mutants. Although a *pif6*<sup>-/-/+</sup> line was generated, testing whether re-expression of the *PIF6* ORF rescues the DNA damage phenotype of *pif6*<sup>-/-</sup> cells was unsuccessful. A clonal survival assay testing the MMS sensitivity of *pif6*<sup>-/-/+</sup> cells produced anomalous results and further testing using the HU sensitivity of *pif6*<sup>-/-/+</sup> cells may be more appropriate as *pif6* mutants display a stronger HU survival phenotype.

Immunolocalisation of PIF6 in BSF cells, using a PIF6 12myc tagged cell line, showed localisation to the nucleus by fluorescence microscopy. This is consistent with Liu *et al.* (2009b) who reported that PIF6 localises to the nucleus of PCF cells. Unlike 12myc RECQ2 and MUS81 12myc, PIF6 12myc did not form foci following phleomycin damage, though nuclear PIF6 12myc fluorescence did increase in intensity. This nuclear localisation did not appear to be cell cycle dependent. An attempt to test the functionality of PIF6 12myc by assaying the MMS sensitivity of PIF6 12myc/- cells produced inconclusive data and it was not possible to conclude whether PIF6 12myc is functional. Repeating this assay using HU, which produces a more distinct phenotype in *pif6* mutants, may reveal whether PIF6 12myc is functional.

TbPIF6 is the only one of the eight *T. brucei* Pif1 helicases to have a nuclear function. TbPIF1, 2, 5 and 8 are involved in kinetoplast replication (Liu *et al.*, 2009a; Liu *et al.*, 2009b; Liu *et al.*, 2010a; Wang *et al.*, 2012). Other eukaryotic Pif1 helicases have mitochondrial roles, such as *S. cerevisiae* ScPif1, which functions in repair, maintenance and recombination of mitochondrial DNA (Cheng *et al.*, 2007; Lahaye *et al.*, 1991; O'Rourke *et al.*, 2002). Pif1 helicases perform a wide range of nuclear functions, participating in almost all aspects of genomic DNA metabolism (reviewed in (Bochman *et al.*, 2010)). However, the phenotype of *T. brucei* mutants described here, particularly increased resistance to MMS, does not appear similar to any Pif1 helicase mutant described to date.

PIF6 appears to play a distinct role in *T. brucei* DNA damage repair. Unlike *mus81* and *recq2* mutants, *pif6*<sup>+/-</sup> mutants display a phenotype, indicating that loss of one PIF6 allele is sufficient to result in a change in the response to HU and MMS damage. Also in contrast to *mus81* and *recq2* mutants, *pif6* mutants displayed an enhanced ability to survive MMS damage. *pif6* mutant sensitivity to

HU is consistent with the increased sensitivity of *S. pombe pfh1* (a Pif1 helicase) mutants to HU (Pinter *et al.*, 2008). More unusual is the MMS resistance of *pif6* mutants – the opposite phenotype to the MMS sensitivity exhibited by *S. cerevisiae pif1Δ* mutants (Budd *et al.*, 2006). A search of the available literature failed to return any reports of resistance to a DNA damaging agent in a Pif1 helicase mutant, suggesting that the increased MMS resistance of *pif6* mutants may be unique among Pif1 helicases. Though increased MMS resistance following *PIF1* knockout has not been observed in other organisms, deletion of *pif1* suppresses the MMS sensitivity of *dna2* mutants in *S. cerevisiae* (Budd *et al.*, 2006). Increased resistance to DNA damage following knockout of a DNA repair factor is unusual, though has been reported for DNA ligase IV. Loss of DNA ligase IV confers increased resistance to replication-associated DSBs in human cells (Kurosawa *et al.*, 2013). These authors found that repair of replication-associated DSBs by NHEJ, in which DNA ligase IV is crucial, is detrimental to cell survival. In the absence of DNA ligase IV, repair of replication DSBs by HR is not suppressed and HR-mediated repair is used, resulting in increased cell survival. If PIF6 expression leads to use of a repair pathway following MMS damage that is detrimental to cell survival, its loss and the subsequent use of a repair pathway resulting in higher cell survival could explain the increased MMS resistance of *pif6* mutants and is an intriguing possibility requiring further study.

In summary, PIF6 is a nuclear *T. brucei* repair factor with a role in DNA repair distinct to the two other factors described in this chapter, *RECQ2* and *MUS81*. Its loss confers increased resistance to MMS, a phenotype not previously reported for a Pif1 helicase. The exact role PIF6 plays in DNA repair and how loss of even a single allele leads to an increased ability to survive MMS damage, while conferring sensitivity to HU, will require further analysis. Testing for rescue of the DNA damage phenotype of *pif6* mutants is an important future experiment, along with further investigation of the MMS sensitivity of mutants, due to variability observed in experiments over time. Examination of the repair products of *pif6* mutants following damage and analysis of the integrity of the nuclear genome of *pif6* mutants would be useful first steps in understanding the role PIF6 plays in DNA repair.

## **Chapter 4**

### **Analysis of the role in double strand break (DSB) DNA repair and VSG switching of RECQ2, PIF6 and MUS81**

## 4 Analysis of the role in double strand break (DSB) DNA repair and VSG switching of RECQ2, PIF6 and MUS81

### 4.1 Introduction

In the previous chapter *RECQ2*, *MUS81* and *PIF6* heterozygous (+/-) and double knockout (-/-) mutants were generated in BSF *T. brucei* and their *in vitro* growth rate and sensitivity to the DNA damaging agents hydroxyurea, phleomycin and MMS of these mutants was analysed. This revealed that *recq2* and *pif6* mutants, both heterozygous and homozygous, as well as *mus81*<sup>-/-</sup> mutants have a growth defect. For *RECQ2*, it was shown that both +/- and -/- mutants have a striking increase in MMS sensitivity, and greater sensitivity to phleomycin and hydroxyurea (HU) was observed in *recq2*<sup>-/-</sup> mutants. *mus81*<sup>-/-</sup> cells were more sensitive to both phleomycin and MMS, whilst *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> mutants displayed similar phenotypes to one another: both were more resistant to MMS and mildly more sensitive to HU. Finally, immunolocalisation of the three proteins endogenously tagged with myc was described, which suggested that *RECQ2* and *MUS81* are involved in the response to phleomycin-induced DNA damage since both were shown to localise in damage-induced foci, with some co-localisation with the RAD51 recombinase.

These data suggest that all three proteins are involved in the response to DNA damage, though act in diverse ways. However, two of the damaging agents used, MMS and phleomycin, result in widespread DNA damage of a number of different types. For example, phleomycin causes DSBs and SSBs (Reiter *et al.*, 1972). The effects of MMS are diverse, including base alkylation, replication fork stalling due to dNTP pool depletion and DNA DSBs that may arise indirectly as a result of DNA alkylation (Bianchi *et al.*, 1986; Brookes & Lawley, 1961; Lundin *et al.*, 2005; Wyatt & Pittman, 2006). The lack of specificity of damage makes it difficult to ascertain the type(s) of damage to which the proteins normally react to and how they function in repair. This chapter will therefore next look specifically at whether *RECQ2*, *MUS81* and *PIF6* are involved in DSB repair.

It was chosen to look specifically at DSB repair due to the proposed role of DSBs in the current model for the initiation of VSG switching. VSG switching by recombination necessitates that free DNA ends are available for recombination, thus requiring a DNA break to be formed to initiate switching. The current hypothesis for the nature of this break is that DSBs occur in the active expression site, specifically in the region of the 70 bp repeats, initiating VSG recombination. DSBs, or at least DNA breaks, have been detected by a ligation-mediated PCR assay in the active VSG BES (Boothroyd *et al.*, 2009; Glover *et al.*, 2013a) around the 70 bp repeats and telomeric VSG; such breaks appear to increase in abundance when telomere integrity is impaired (Jehi *et al.*, 2014b). However, whether the putative DSBs are limited to the active expression site, as originally reported (Boothroyd *et al.*, 2009) is unclear, since later studies did not detect this association with transcription status (Glover *et al.*, 2013a). Nonetheless, both Boothroyd *et al.* (2009) and Glover *et al.* (2013a) utilised the inducible expression of a yeast meganuclease (I-SceI) to show that when DSBs are introduced into the active (but not inactive) expression site, the rate of VSG switching increases.

This chapter describes the generation of *recq2*, *mus81* and *pif6* mutants in cell lines that allow the targeted investigation of the role of these factors in DSB repair and in VSG switching. To investigate their role in DSB repair, I-SceI cell lines were used as in Glover *et al.* (2008, 2011 & 2013a), in which it is possible to induce a single, temporarily and spatially controlled DSB at the active VSG BES or at a chromosome internal location. For examination of the proteins' role in VSG switching, this chapter also describes the generation of mutants in a cell line that uses a thymidine kinase negative selection system, essentially as used previously (Povelones *et al.*, 2012), to assay VSG switching rate and mechanism *in vitro*. By examining the role of RECQ2, MUS81 and PIF6 in the repair of induced DSBs and comparing these data with that of their role in VSG switching, this chapter aims to build a more detailed picture of how these proteins function in DNA repair and VSG switching in *T. brucei*. This chapter will also explore the contribution of DSB repair to VSG switching. In examining DSB repair and VSG switching, the work tests the hypothesis that DSBs act as the initiating lesion for VSG switching.

## 4.2 Description of I-SceI cell lines used

In order to investigate the role of RECQ2, MUS81 and PIF6 in DSB repair, we utilised two cell lines in which a single DSB could be induced, HR1 and HRES (Gift, David Horn). The HR1 and HRES cell lines (Fig. 4-1) utilise the I-SceI meganuclease to cleave DNA to form a DSB. I-SceI is a meganuclease encoded by a mitochondrial group I intron of *S. cerevisiae* that cleaves a specific 18 bp recognition sequence (Colleaux *et al.*, 1988); this sequence does not appear naturally in the *T. brucei* genome. Therefore, in *T. brucei* in which the I-SceI recognition sequence has been introduced, induction of I-SceI expression results in a single DSB at the I-SceI recognition site.

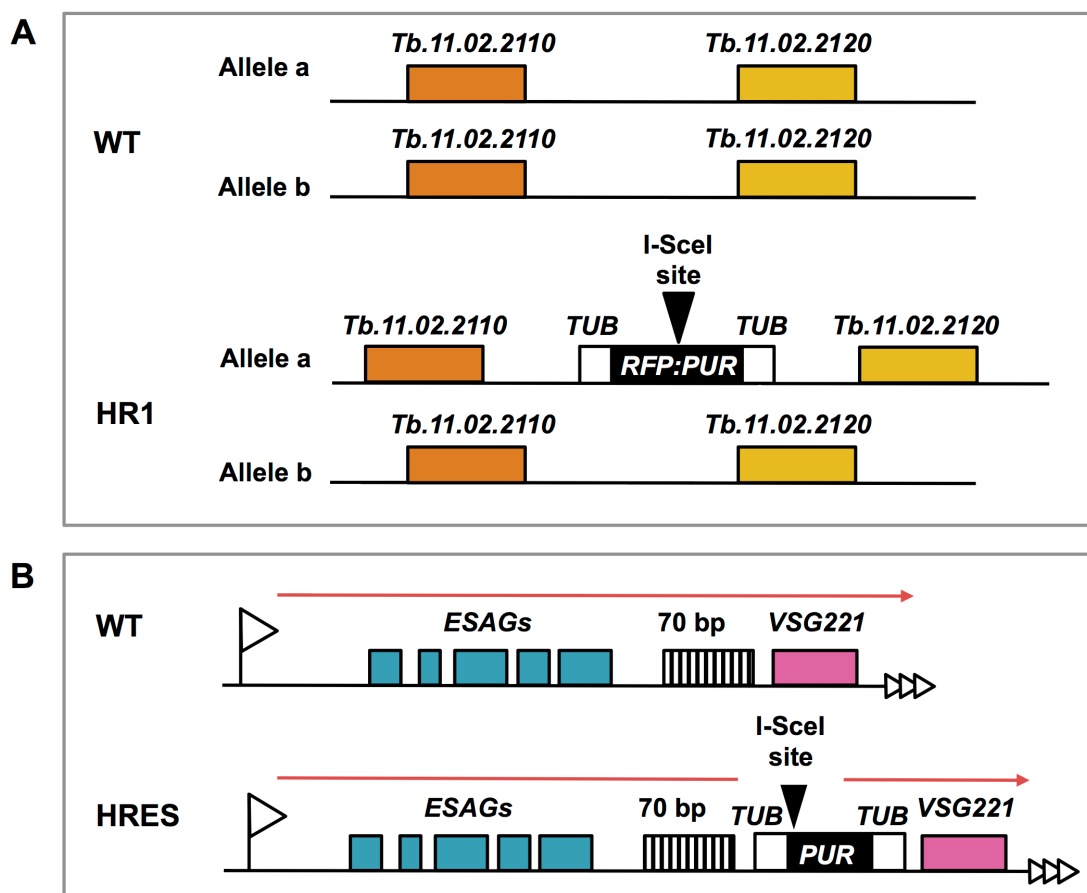
Both HR1 and HRES cell lines contain the I-SceI recognition site, the I-SceI ORF under the control of a tet-on promoter and the tet repressor protein (TetR). The TetR is expressed from the tubulin locus and is maintained by the presence of the antibiotic resistance gene bleomycin (*BLE*) (Glover *et al.*, 2007). The I-SceI ORF is fused to an N-terminal SV40 nuclear localisation signal expressed from an RRNA spacer locus and maintained by the presence of the antibiotic resistance gene hygromycin (Glover *et al.*, 2007). This arrangement of I-SceI expression control by the TetR means that expression of I-SceI is induced by addition of tetracycline to the medium. In the absence of tetracycline, I-SceI expression is repressed by the binding of the TetR to the tet operator. Upon addition of tetracycline to the culture medium, tetracycline binds to the TetR, preventing it binding to the tet operator. Repression of I-SceI expression is then relieved and I-SceI protein recognises and cleaves the I-SceI recognition sequence. DSB induction is therefore controlled both temporarily by tetracycline, and spatially by a single I-SceI recognition site location (below).

The I-SceI recognition site differs between the two cell lines. In HR1, the I-SceI recognition site is located on chromosome 11 between genes Tb.11.02.2110 and Tb.11.02.2120 (genedb.org), and >1 Mbp from the nearest telomere (Glover *et al.*, 2008). Here, the I-SceI site is embedded within an *RFP* (red fluorescent protein):*PUR* (puromycin N-acetyl transferase) fusion gene (Glover *et al.*, 2008). The *RFP:PUR* gene is flanked by tubulin sequences for mRNA transplicing and polyadenylation. Thus, HR-directed repair after I-SceI-induced DSB formation could occur by recombination between chromosome 11a (containing the I-SceI

site) and its homologue (11b), but could also occur ectopically with chromosome 1 (where the tubulin locus is found) using the short tubulin sequences on the *RFP:PUR* cassette. Despite the presence of the *RFP* gene, Glover *et al.* (2008) reported that no fluorescence is detectable in HR1 cells. We were also unable to detect any fluorescence in HR1 cells (data not shown). In HRES, the I-SceI recognition site is located upstream of *VSG221* in the active *VSG* BES on chromosome 6a, fused to a *PUR* gene (Glover *et al.*, 2013a). Here DSB induction has been proposed to mimic *VSG* switching by initiating HR through available homology (e.g. other *VSGs*, 70 bp repeats, telomere repeats, *ESAGs*). In both HR1 and HRES, *PUR* gene presence at the I-SceI recognition site provides a means to assay for I-SceI cutting in induced cell cultures. DNA end resection following I-SceI cutting results in puromycin sensitivity due to the proximity of the *PUR* gene to the I-SceI target sequence, meaning it must be degraded to access the flanking homology that drives DSB repair (Glover *et al.*, 2008).

I-SceI induction in the HR1 cell line allows the investigation of DSB repair at a chromosome-internal location. In contrast, the HRES cell line enables study of DSB repair at a telomere and active expression site. Both Boothroyd *et al.* (2009) and Glover *et al.* (2008 & 2013a) demonstrated, by Southern blot analysis, that I-SceI induction results in the formation of a DSB at the I-SceI recognition site. Specifically, the I-SceI cleavage site is located in the region in which DNA breaks triggering *VSG* switching are hypothesised to occur and the experimental setup in the HRES is very similar to that described by Boothroyd *et al.* (2009). Both the HR1 and HRES cell lines have been used previously to investigate DSB repair in *T. brucei*; Glover *et al.* (2008) generated and used HR1 to investigate DSB repair, revealing the predicted HR events (above), as well as MMEJ, where repair normally occurs intrachromosomally based on very short flanking sequences and independent of HR (e.g. in the absence of RAD51) (Burton *et al.*, 2007; Conway *et al.*, 2002b; Glover *et al.*, 2011; Glover *et al.*, 2008). The HRES line was used by Glover *et al.* (2013a) (there referred to as *VSG<sup>up</sup>*), where they found that I-SceI induced breaks at the 70 bp repeats induced *VSG* switching, consistent with the findings of Boothroyd *et al.* (2009).





**Figure 4-1** I-SceI target sequence organisation in HR1 and HRES cell lines  
As described in the text: (A) the HR1 cell line contains an I-SceI recognition site embedded within an *RFP:PUR* fusion gene (black), flanked by tubulin sequences (white), located between genes *Tb.11.02.2110* and *Tb.11.02.2120* on chromosome 11a; (B) the HRES cell line contains an I-SceI recognition site upstream of a *PUR* gene, flanked by tubulin sequences, located downstream of the 70 bp repeats of the active *VSG221* expression site on chromosome 6.

## 4.3 Generation of knockout mutants in I-SceI lines

### 4.3.1 Generation of knockout lines

It was attempted to generate heterozygous (+/-) and double knockout (-/-) mutants in HR1 and HRES I-SceI cell lines following the same procedure as used in the wild type BSF cells (described in Section 3.3.1). Detailed explanation for the knockout strategy and the cloning procedure of the knockout plasmids is in Section 3.3.1.

HR1 and HRES cell lines were first transformed with  $\Delta RECQ2::BSD$ ,  $\Delta MUS81::BSD$  or  $\Delta PIF6::BSD$  in an attempt to generate +/- cell lines. In all cases, antibiotic resistant clones were selected using  $10 \mu\text{g.mL}^{-1}$  blasticidin as well as  $0.2 \mu\text{g.mL}^{-1}$  puromycin,  $5 \mu\text{g.mL}^{-1}$  hygromycin and  $2 \mu\text{g.mL}^{-1}$  phleomycin to maintain the

I-SceI meganuclease recognition site, I-SceI meganuclease and tetracycline repressor constructs, respectively.

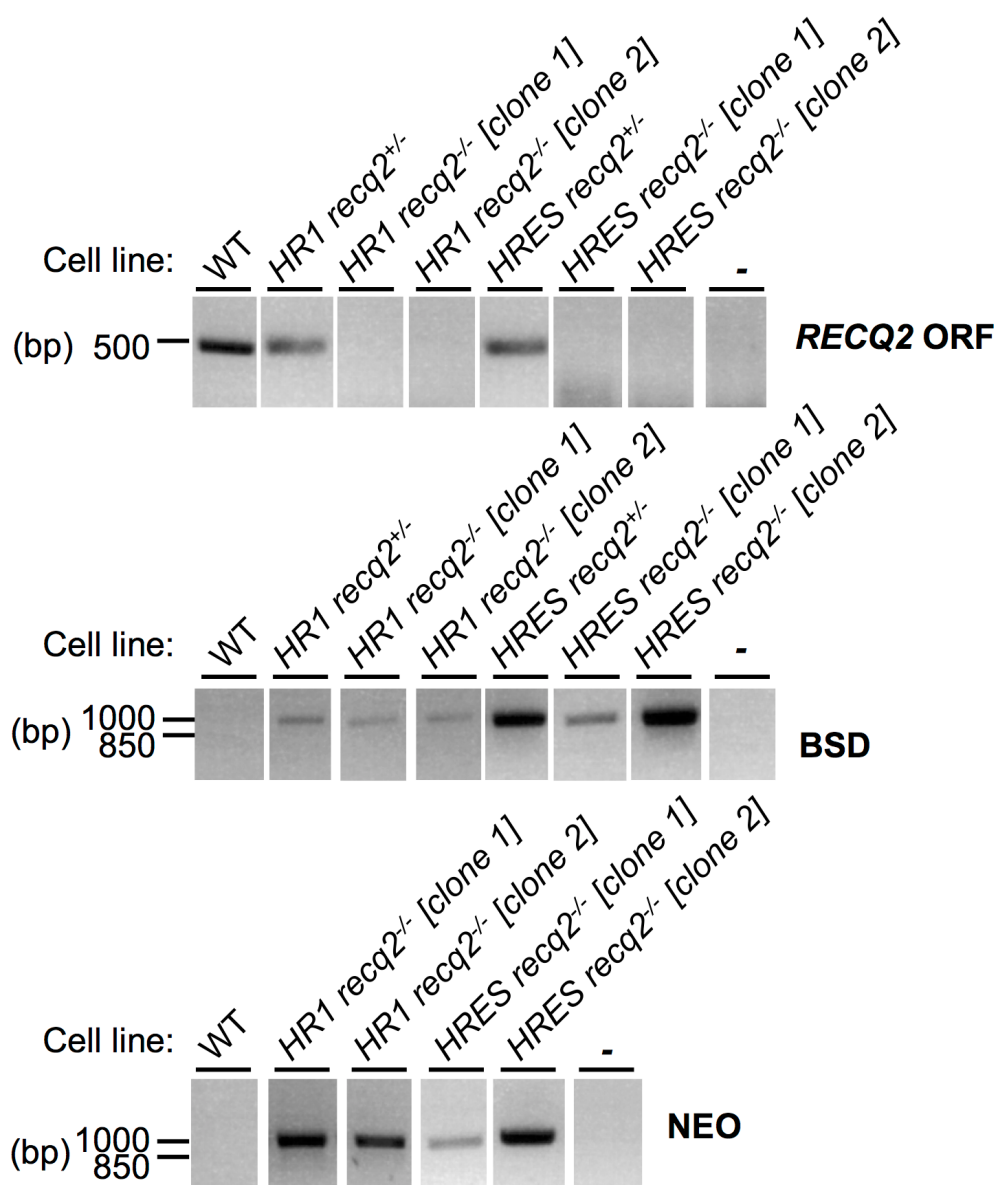
PCR analysis was used to confirm the generation of heterozygous mutants (as described previously in wild type cells, see Section 3.3.3) using genomic DNA extracted from drug resistant transformant clones. Six putative *HR1 recq2*<sup>+/-</sup> clones and six *HRES recq2*<sup>+/-</sup> clones were tested, all of which were correct. Six putative *HR1 mus81*<sup>+/-</sup> clones were tested, of which five were correct. Five putative *HRES mus81*<sup>+/-</sup> clones were tested, all of which were correct. Six putative *HR1 pif6*<sup>+/-</sup> clones and six *HRES pif6*<sup>+/-</sup> clones were tested, all of which were correct. (In all cases, correct clones were defined by the blasticidin knockout construct having integrated as expected; data not shown, PCR described below.) One confirmed +/- mutant from each of the  $\Delta RECQ2::BSD$ ,  $\Delta MUS81::BSD$  and  $\Delta PIF6::BSD$  transformants was chosen and transformed with the  $\Delta RECQ2::NEO$ ,  $\Delta MUS81::NEO$  or  $\Delta PIF6::NEO$  construct, respectively. Transformants were selected with 5  $\mu\text{g.mL}^{-1}$  blasticidin and 2.5  $\mu\text{g.mL}^{-1}$  G418, in addition to 0.2  $\mu\text{g.mL}^{-1}$  puromycin, 5  $\mu\text{g.mL}^{-1}$  hygromycin and 2  $\mu\text{g.mL}^{-1}$  phleomycin (as in the first round of transformation). A number of the clones recovered from the second round of transformation were tested by PCR (see below) to confirm knockout of the target gene. Five putative *HR1 recq2*<sup>-/-</sup> clones were tested, of which four were correct, and three *HRES recq2*<sup>-/-</sup> clones were tested, of which two were correct and are shown in Figure 4-2. Six putative *HR1 mus81*<sup>-/-</sup> clones were tested, of which all were correct, and four *HRES mus81*<sup>-/-</sup> clones were tested, of which two were correct and are shown in Figure 4-3. Eight putative *HR1 pif6*<sup>-/-</sup> clones were tested, of which all were correct, and ten *HRES pif6*<sup>-/-</sup> clones were tested, of which eight were correct and two are shown in Figure 4-4.

### 4.3.2 Confirmation of knockout mutants by PCR

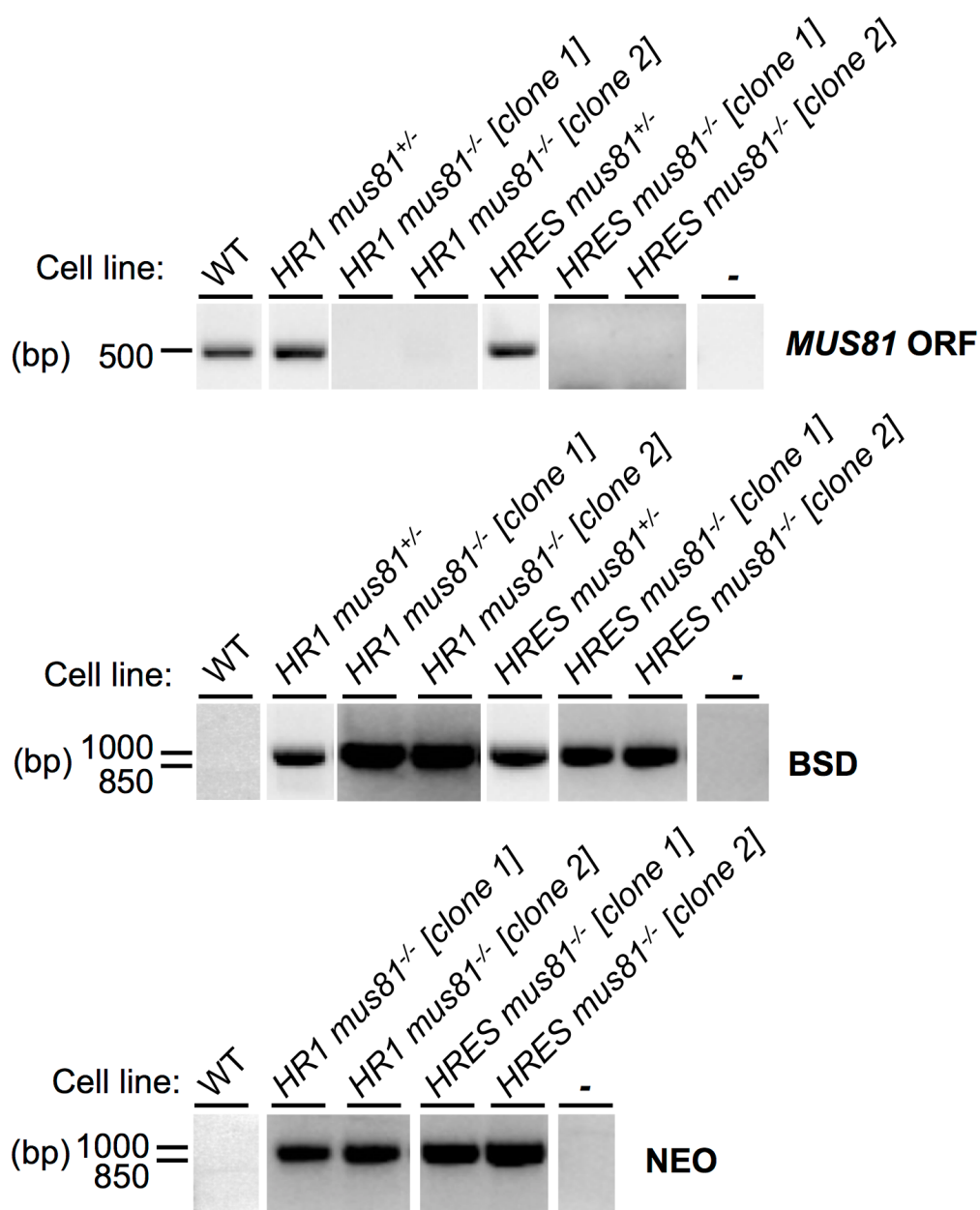
Demonstration of the generation of heterozygous (+/-) and homozygous (-/-) mutants of each gene in both the HR1 and HRES cells was confirmed by a number of PCRs. Figures 4-2, 4-3 and 4-4 show PCRs for the clones that were selected for further analysis. Three different PCRs were performed (described in Section 3.3.3) to confirm integration of the knockout cassettes in each cell line. A region of the ORF was PCR-amplified ("ORF PCR", (*RECQ2* primers #44 & #45;

*MUS81* primers #46 & #47; *PIF6* primers #48 & #49)) to determine if the gene had been deleted. In addition, a ~900-1300 bp region was PCR-amplified using a forward primer lying upstream of the 5' UTR region present in the knockout construct (*RECQ2* primer #51, *MUS81* primer #52, *PIF6* primer #53) and a reverse primer specific to the BSD or NEO gene ("BSD PCR" (primer #154) and "NEO PCR" (primer #155), respectively). These resistance gene PCRs were designed to confirm the presence and correct integration location of the knockout cassette relative to the target gene. Gene-specific PCR details are listed in the legend of Figures 4-2, 4-3 and 4-4 and sequences of all primers can be found in Appendix 7.1.

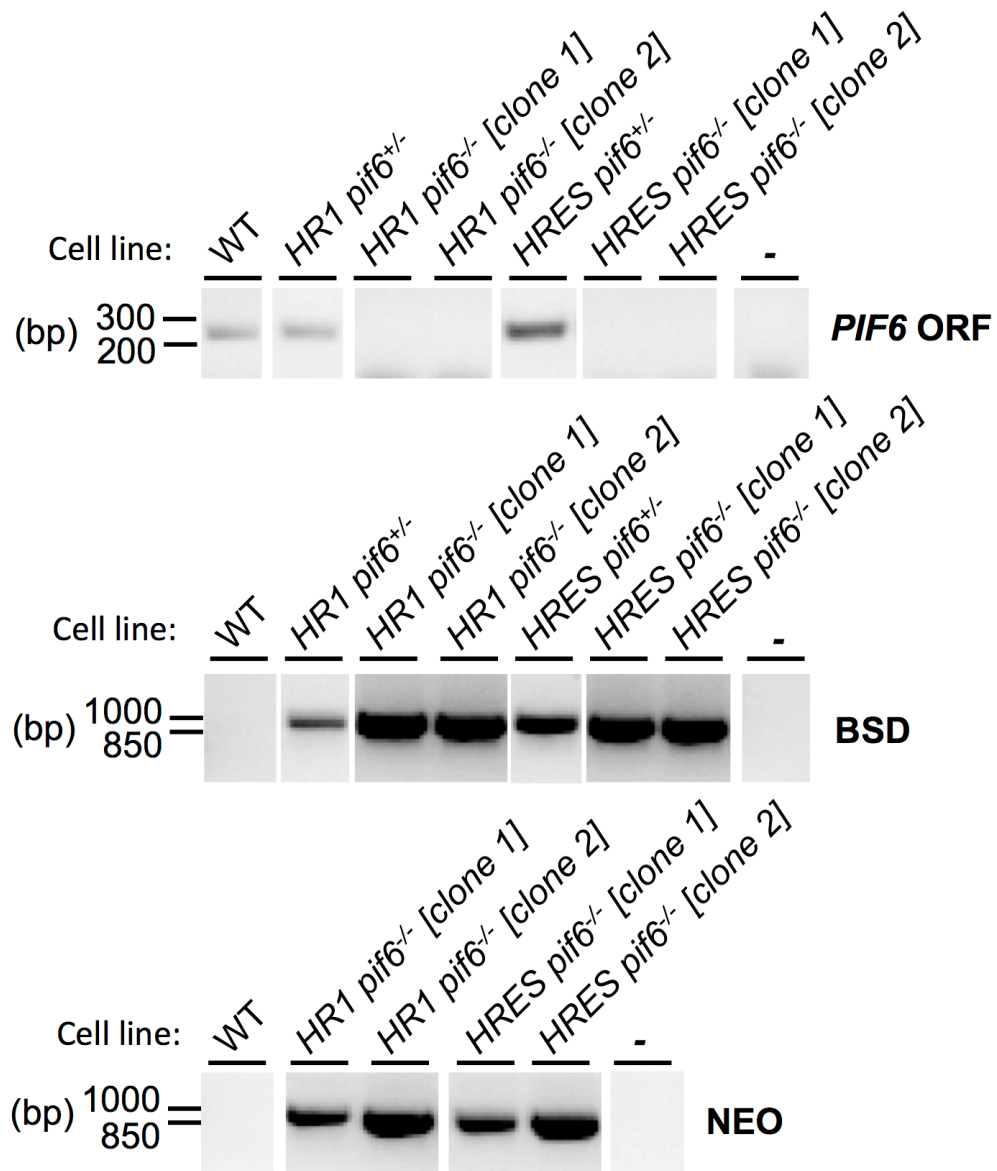
The agarose gels in Figures 4-2, 4-3 and 4-4 show confirmation, using the PCRs described above, of the correct integration of the knockout cassettes in the HR1 and HRES cell lines and absence of the cognate ORF in the different -/- clones. In all cases, two -/- mutants were examined in subsequent experiments to test for reproducibility of data. Only the expected knockout cassette could be PCR-amplified ("BSD PCR") in each heterozygote line, whereas both cassettes could be amplified in homozygous knockout cell lines. Furthermore, the ORF of the targeted gene could no longer be PCR-amplified in the double knockout cell lines, indicating successful deletion of each gene in HR1 and HRES backgrounds.



**Figure 4-2** PCR confirmation of *RECQ2* knockout in HR1 and HRES cell lines  
 Agarose gels of PCR products generated from genomic DNA using primers described in the text. A 498 bp region of the open reading frame (primers #44 & #45) was PCR amplified (“RECQ2 ORF”). Integration of the blasticidin knockout construct was tested by PCR amplification using primers #51 & #154 (1000 bp PCR product) that bind in the *BSD* resistance cassette and upstream of *RECQ2* (“BSD”). Integration of the *NEO* knockout construct was tested by PCR amplification using primers #51 and #155 (1008 bp PCR product) that bind in the *NEO* resistance gene and upstream of *RECQ2* (“NEO”). Primer sequences are listed in Appendix 7.1. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels or from disparate parts of the same gel; size markers are shown (ladder, bp).



**Figure 4-3 PCR confirmation of *MUS81* knockout in HR1 and HRES cell lines**  
 Agarose gels of PCR products generated from genomic DNA using primers described in the text. A 494 bp region of the *MUS81* open reading frame (primers #46 & #47) was PCR amplified ("MUS81 ORF"). Integration of the blasticidin knockout construct was tested by PCR amplification using primers #52 & #154 (966 bp PCR product) that bind in the blasticidin resistance cassette and upstream of *MUS81* ("BSD"). Integration of the *NEO* knockout construct was tested by PCR amplification using primers #52 and #155 (974 bp PCR product) that bind in the *NEO* resistance gene and upstream of *MUS81* ("NEO"). Gaps indicate that lanes have been aligned in this figure after excision from multiple gels or from disparate parts of the same gel; size markers are shown (ladder, bp).



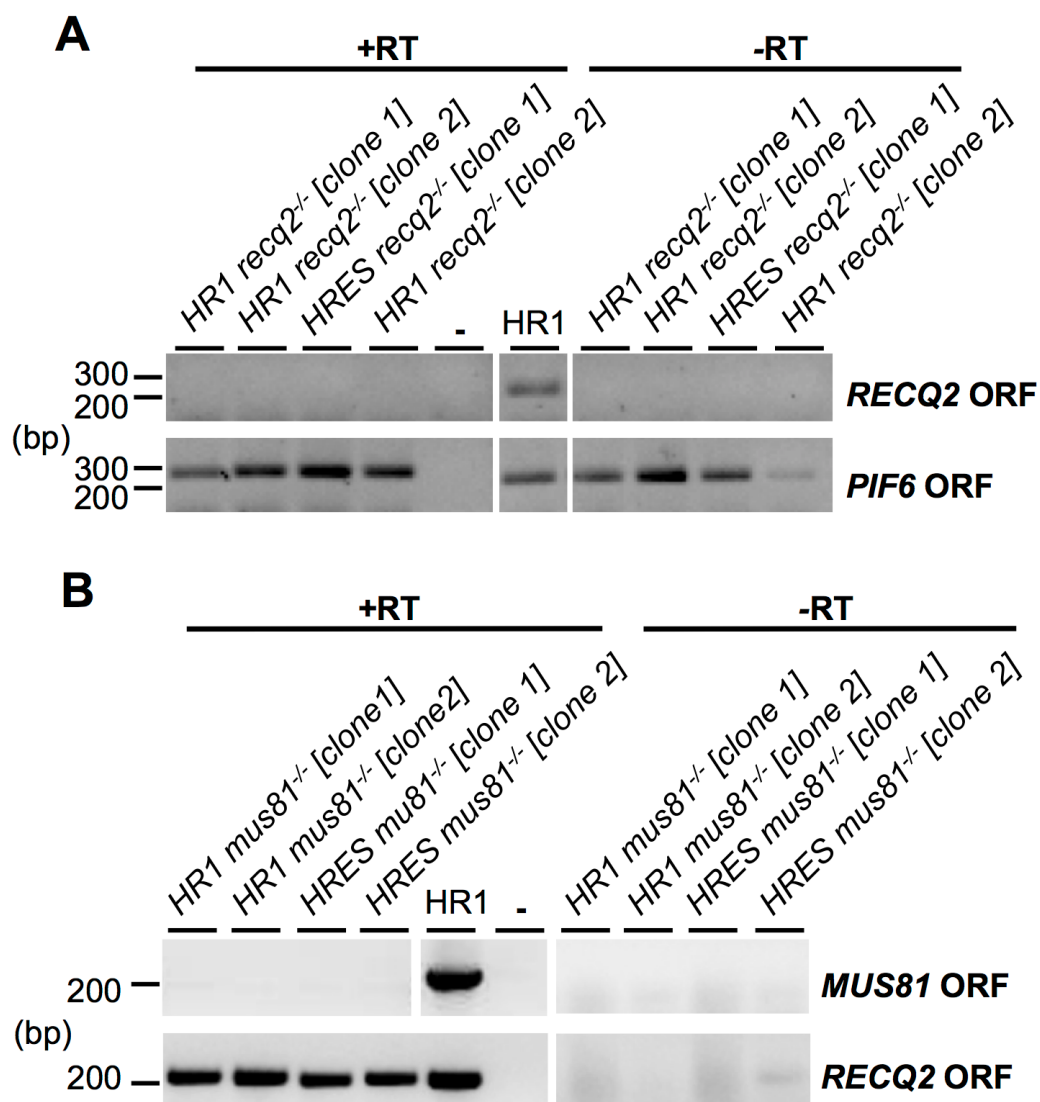
**Figure 4-4** Confirmation by PCR of *PIF6* knockout in HR1 and HRES cell lines  
Agarose gels of PCR products generated from genomic DNA using primers described in the text. A 245 bp region of the *PIF6* open reading frame (primers #81 & #82) was PCR amplified ("PIF6 ORF"). Integration of the blasticidin knockout construct was tested by PCR amplification using primers #53 & #154 (890 bp PCR product) that bind in the blasticidin resistance cassette and upstream of *PIF6* ("BSD"). Integration of the blasticidin knockout construct was tested by PCR amplification using primers #53 & #155 (898 bp PCR product) that bind in the *NEO* resistance gene and upstream of *PIF6* ("NEO"). Primer sequences can be found in Appendix 7.1. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels or from disparate parts of the same gel; size markers are shown (ladder, bp).

### 4.3.3 Confirmation of *recq2* and *mus81* knockout lines by RT-PCR

RT-PCR was used to confirm absence of expression of the deleted gene in each knockout (-/-) cell line. A ~200 bp region of the targeted gene ORF was amplified from cDNA, as described in Section 3.3.4, using primers #77 & #78 (*RECQ2*), #79 & #80 (*MUS81*) and #81 & #82 (*PIF6*). In addition, PCR

amplification of a ~200 bp region of another gene (*RECQ2*, *MUS81* or *PIF6*) that was not targeted, was used as a positive control to confirm the presence of cDNA in the sample. Sequences of all primers can be found in Appendix 7.1 and gene-specific PCR details are described in the legend of Figure 4-5.

The agarose gel in Figure 4-5 shows that the segment of the ORF of *RECQ2* and *MUS81* could not be PCR amplified in their respective knockout (-/-) cell lines, but that the control gene could be PCR amplified in each case. Genomic DNA contamination was present in the *RECQ2* samples, as indicated by PCR amplification of the control (*PIF6*) gene ORF in some -RT cDNA samples. However, the *RECCQ2* ORF PCR only generated a product in untransformed HR1 cells, which acted as a control, and generated no product in all the double knockout cell lines analysis in both the +RT and -RT preparations. Thus, despite some contamination, expression of the both targeted genes in each -/- cell line could no longer be detected and knockout of these genes was successful in both HR1 and HRES cell line backgrounds.



**Figure 4-5** Confirmation by RT-PCR of *recq2* and *mus81* knockout cells in HR1 and HRES cell lines

cDNA was synthesised from RNA extracted from HR1 and HRES wild type and *recq2* and *mus81* HR1 and HRES mutants. (A) RT-PCR check of *recq2* mutants: A 325 bp region of the *RECQ2* ORF (primers #77/#78) and 245 bp region of the *PIF6* ORF (primers #81/#82) was PCR amplified from cDNA of HR1 and HRES wild type cells and *recq2* mutants (+RT), as well as from samples in which no reverse transcriptase had been added (-RT). (B) A 214 bp region of the *MUS81* ORF (primers #79/#80) and a 214 bp region of the *RECQ2* ORF (as described in (A)) was PCR amplified from cDNA of HR1 and HRES wild type cells and *mus81* mutants (+RT), as well as from samples in which no reverse transcriptase had been added (-RT). Distilled water was used as a negative control. Gaps indicate different gels or portions of gels aligned in this figure. Sizes shown, ladder (bp).

#### 4.4 Analysis of DSB repair at a chromosome internal location

In order to understand if any of the gene knockouts had an effect on cell survival following induction of a chromosome-internal DSB, the clonal survival of I-SceI-

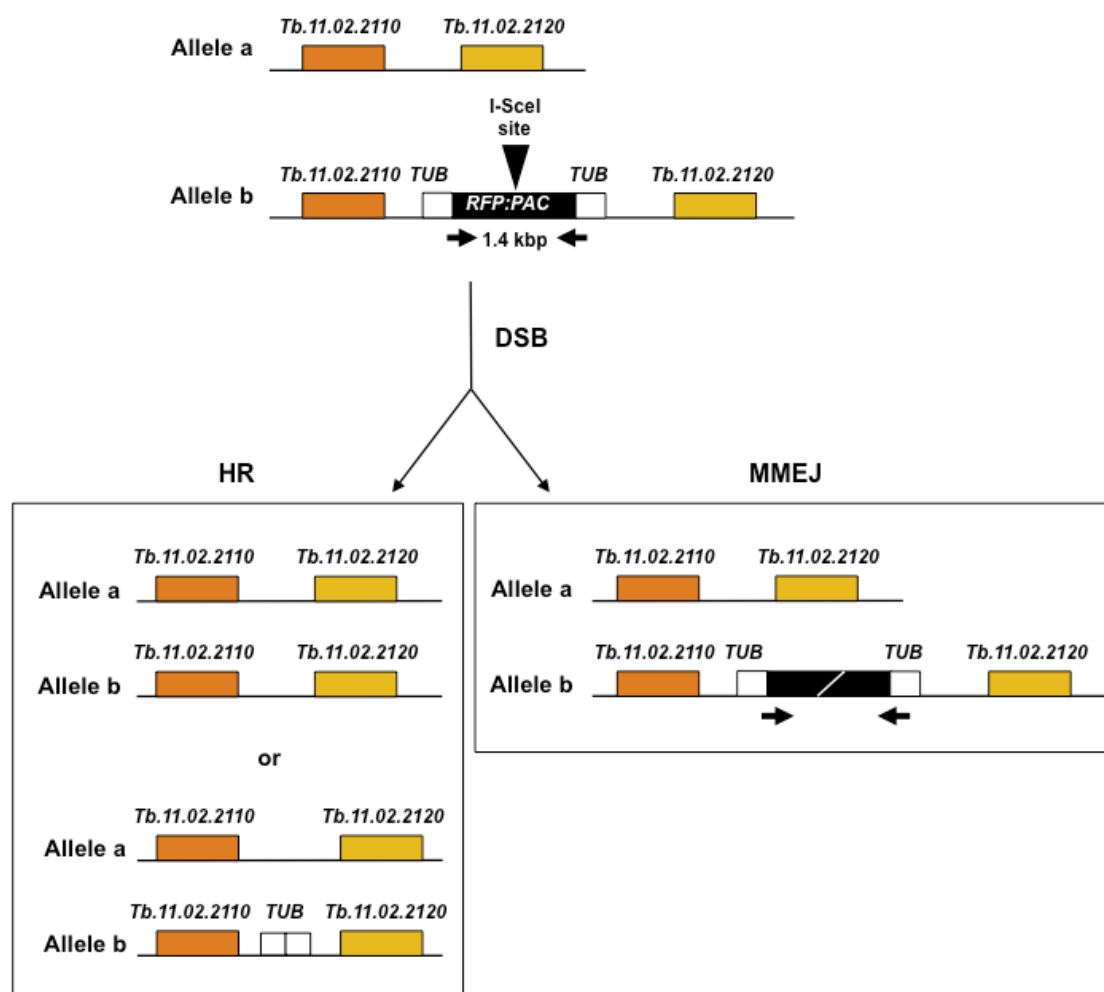


induced *recq2*, *mus81* and *pif6* mutants in an HR1 background was assayed. To survive, cells in which a DSB has been induced must repair the break. Changes in the clonal survival following DSB induction in mutants can therefore reveal whether the knocked out gene is involved in DSB repair.

The clonal survival assay was performed as in Glover *et al.* (2013a) (see Materials and Methods, Section 2.2.3.3). Cells were cultured to mid-log phase ( $1 \times 10^6$  cells.mL<sup>-1</sup>) in the presence of phleomycin, puromycin and hygromycin to maintain the I-SceI genetic components in the cells prior to the assay. Additionally, the cells were grown in tetracycline-free HMI-9 to limit induction of I-SceI expression. Mutant and non-mutant HR1 cultures were distributed at a sub-clonal dilution (0.26 cells per well) over four 96 well plates either with or without I-SceI induction. I-SceI expression was induced by addition of tetracycline (Calbiochem) to the cultures to a final concentration of 2 µg.mL<sup>-1</sup>. Plates were incubated for seven to ten days and scored for survival (normalising to un-induced cultures). HMI-9 medium, which in the absence of growing cells is red, becomes orange/yellow when saturated with live cells. Therefore orange/yellow wells were scored as positive. When there was doubt using this method, wells were examined microscopically to determine whether they contained live cells.

For each cell line six I-SceI uninduced and up to 25 I-SceI induced surviving clones were selected for analysis. Successful I-SceI induction and repair of the subsequent DSB should result in the puromycin resistance gene adjacent to the I-SceI site being deleted or truncated and thus surviving clones should be puromycin sensitive (Glover *et al.*, 2008). Therefore, clones taken for analysis were tested for puromycin resistance, using 1 µg.mL<sup>-1</sup>. In addition, genomic DNA was extracted for PCR analysis to attempt to determine whether the induced DSB was repaired by HR or MMEJ. Figure 4-6 shows this PCR. Primers #86 and #87 were used (Glover *et al.*, 2011), which PCR amplify a 1.4 kbp region of the *RFP:PUR* cassette. Extensive DNA end resection that occurs in HR results in the loss of the primer binding sites. Therefore, following repair of the I-SceI-induced DSB, clones that had undergone repair by HR would not retain the primer binding sites and the PCR would fail in these clones. Conversely, clones that had repaired the DSB by MMEJ would retain the primer binding sites, due to the requirement of MMEJ for little sequence identity, and assuming that this

microhomology is found within the cassette. In these circumstances, this PCR in MMEJ-repaired clones results in a PCR product smaller than that from uninduced clones (1.4 kbp). Glover *et al.* (2008) have previously shown that approximately 85% of surviving HR1 clones repaired the DSB by HR. This assay allows the comparison of the repair pathway profiles of wild type and *recq2*, *mus81* and *pif6* mutants and provides insight into whether RECQ2 is involved in HR or MMEJ in DSB repair in BSF *T. brucei*.



**Figure 4-6** Repair of I-SceI-induced DSB in HR1 cells

As described in the text, repair of an I-SceI-induced DSB in HR1 cells (allele a) can be repaired via HR using sequence on the homologous chromosome (allele b) or the tubulin sequence on chromosome 1, both resulting in loss of the primer binding sites. Alternatively, repair can proceed via MMEJ, resulting in retention of primer binding sites but an altered PCR product length.

#### 4.4.1 Analysis of chromosome-internal DSB repair in *recq2* mutants

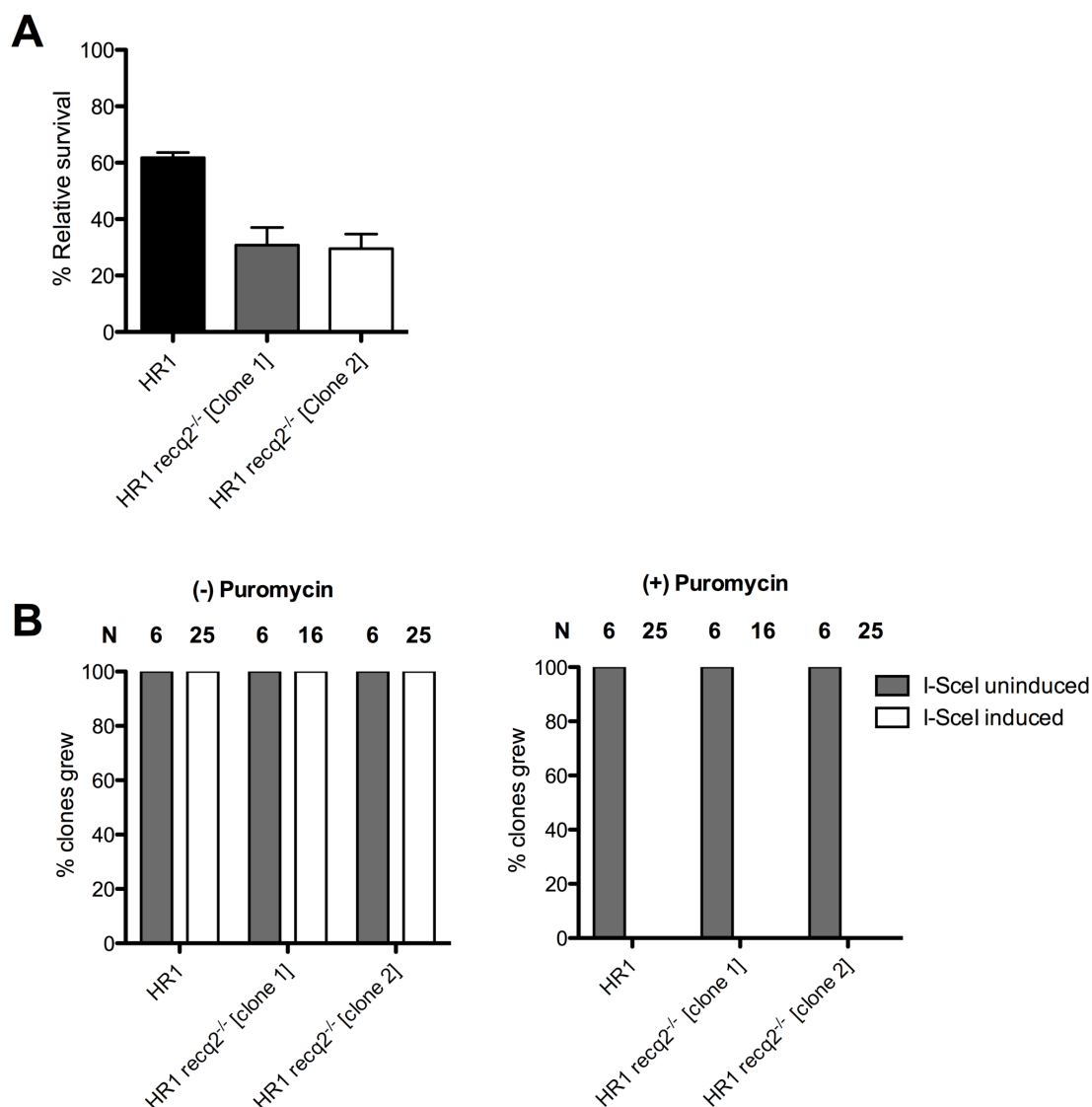
To investigate whether RECQ2 is involved in DSB repair in *T. brucei*, clonal survival assays using *recq2* mutants in the HR1 background were performed

twice. The survival rate of HR1 wild type cells following I-SceI induction was 61.8% (Fig. 4-7A), very similar to the survival rate of ~60% reported by Glover *et al.* (2008). In the absence of RECQ2, the survival rate decreased by 50% compared to HR1 wild type, with survival rates of 30.8% and 29.5% in two *HR1 recq2*<sup>-/-</sup> clones. This indicates that *recq2* mutants are less able to survive an induced DSB, and therefore that RECQ2 is involved in repair of DSBs in BSF *T. brucei*. Previous data showed *recq2* mutant sensitivity to the DNA damaging agents phleomycin and MMS (see Section 3.4.3). When taken together with the finding that *recq2* mutants are impaired in their ability to repair DSBs, these data suggest a wide role for RECQ2 in the repair of DNA damage.

Analysis of the puromycin sensitivity of surviving clones (Fig. 4-7B) showed that all induced survivors were puromycin sensitive and that all un-induced clones were puromycin resistant, indicating that a functional *PUR* gene was lost in all induced clones and confirming successful DSB formation. PCR analysis of clones was performed to attempt to analyse the mechanism of DSB repair in induced clones using the repair PCR described above (see Section 4.4). Despite multiple attempts, PCR amplification was unsuccessful in all but two uninduced clones. This PCR should be successful in uninduced clones, as they have not undergone I-SceI induction and DSB formation. I-SceI induction leads to DSB formation at the I-SceI recognition site and the majority of cells repair this DSB using HR, which results in loss of the primer binding sequences (Fig. 4-6) (Glover *et al.*, 2008). Absence of I-SceI induction in the uninduced clones analysed means that they should all still contain the primer binding sites for this assay. Why this PCR was so inefficient is unclear, but the lack of reproducible amplification in the uninduced controls means the findings are inconclusive and therefore these data are not presented here.

The clonal survival of *HR1 recq2* mutants was assayed a second time, selecting one of the *HR1 recq2*<sup>-/-</sup> mutants and this time assaying a *HR1 recq2*<sup>+/-</sup> mutant, to investigate the effect of loss of a single *RECQ2* allele. The results of this assay (Fig. 4-8A) contrast sharply with the first time the assay was performed. Although the survival of HR1 wild type cells was 43.3%, comparable with that observed previously, the survival of *HR1 recq2*<sup>-/-</sup> cells (48.6%) was higher than HR1 wild type, meaning that the survival rate was substantially higher than that

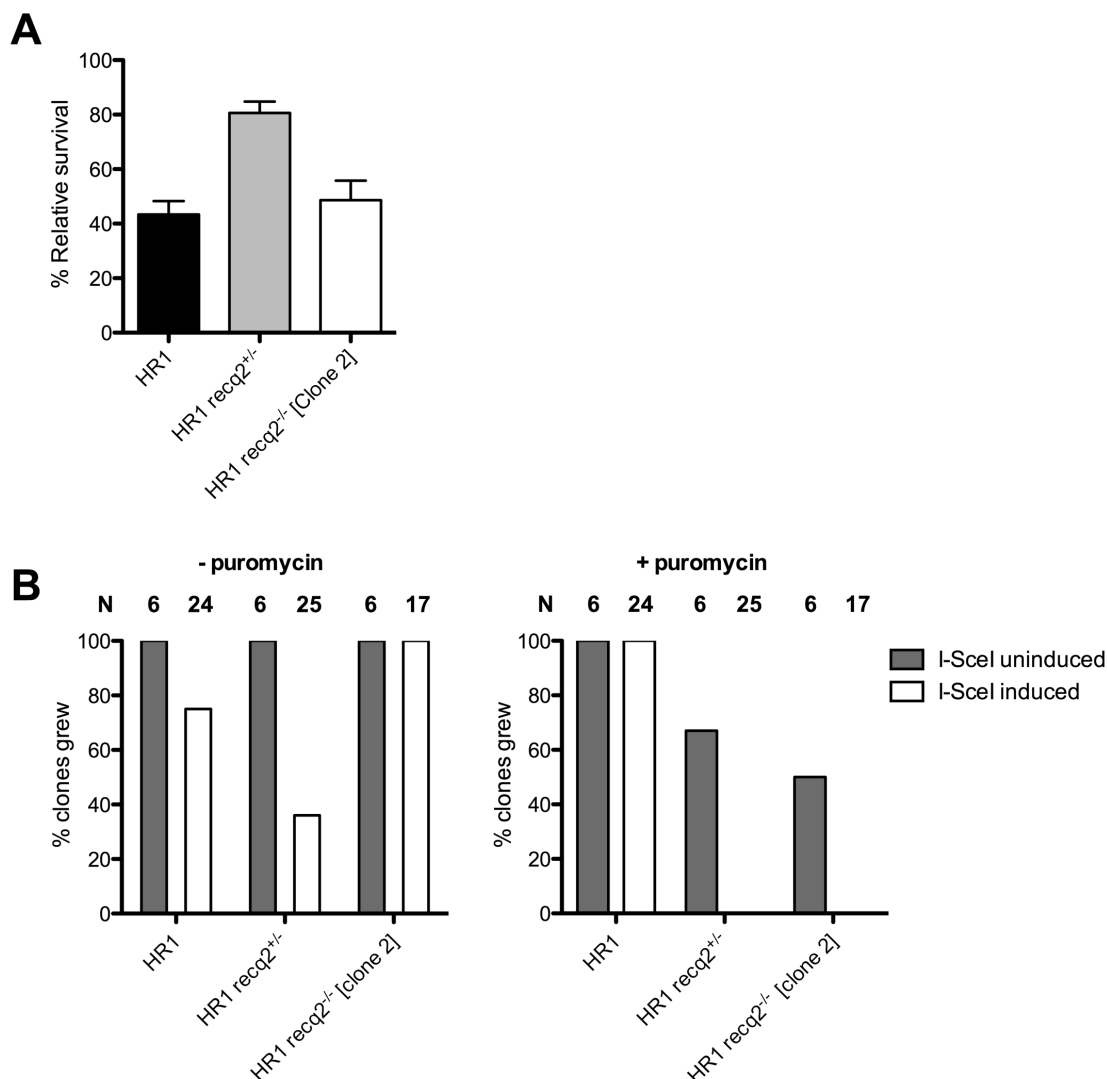
measured for the same clone previously, as well as higher than the other  $-/-$  clone (Fig. 4-7A).



**Figure 4-7 Clonal survival following I-SceI induction in *HR1 recq2* mutants**  
*HR1* wild type and two *HR1 recq2*<sup>-/-</sup> clones were distributed in three 96 well plates at a concentration of 0.26 cells per well either in the absence (I-SceI uninduced) or presence (I-SceI induced) of 2  $\mu\text{g.mL}^{-1}$  tetracycline. The number of wells with surviving cells (clones) was counted after 7-10 days of incubation. (A) Cell survival is shown as the average percentage of clonal survivors following I-SceI induction compared to survivors without I-SceI induction. Error bars represent standard error of the mean (SEM) between the three 96 well plates. (B) Puromycin sensitivity of surviving I-SceI induced and uninduced clones, represented as the percentage of tested clones that grew in the presence (+) or absence (-) of 1  $\mu\text{g.mL}^{-1}$  puromycin. N, number of clones analysed.

There are a number of difficulties in the interpretation of this second experiment. Puromycin sensitivity testing was inconclusive (Fig.4-8B). All uninduced survivors should be puromycin resistant and all induced survivors should be puromycin sensitive. However, only 67% of *HR1 recq2*<sup>+/-</sup> and 50% of

*HR1 recq2<sup>-/-</sup>* uninduced survivors were puromycin resistant. Furthermore, 100% of HR1 wild type induced survivors were puromycin resistant, suggesting that the ‘induced’ HR1 wild type cells were not induced and DSBs did not form. The 40% survival rate of the HR1 wild type cells appears to conflict with this interpretation, however, since it would be expected to be equal to un-induced cultures if DSB induction had not occurred. However, a further difficulty in the interpretation of these data is that not all surviving clones grew well in the absence of puromycin despite the fact that all clones should be expected to survive whether or not they underwent DSB induction. The reason for the differences between the first and second repeats of this experiment is unclear. However, the second experiment was conducted some time after the first and a possible explanation lays in background I-SceI expression leading to instability and uncharacterised changes over time in culture. Though all of the I-SceI cell lines are cultured in ‘tetracycline-free’ medium, leaky expression of I-SceI would result in some background levels of DSB induction, which might cause loss of the I-SceI target sequence over time and puromycin sensitivity in these cells. The nature of the changes selected by this background expression could be variable and this may explain why some clones that survive addition of tetracycline grow poorly even in the absence of puromycin. Taken together, such putative changes may make the survival measurements unreliable. These inconsistencies make this second repeat an unreliable source of data in determining whether RECQ2 plays a role in DSB repair in *T. brucei*. By contrast, in the first experiment it was possible to confirm that all induced cultures underwent DSB induction from the puromycin resistance data, suggesting this experiment provides a more reliable assessment of RECQ2 function in DSB repair.



**Figure 4-8** Repeat of clonal survival assay of HR1 *recq2* mutants following I-SceI DSB induction

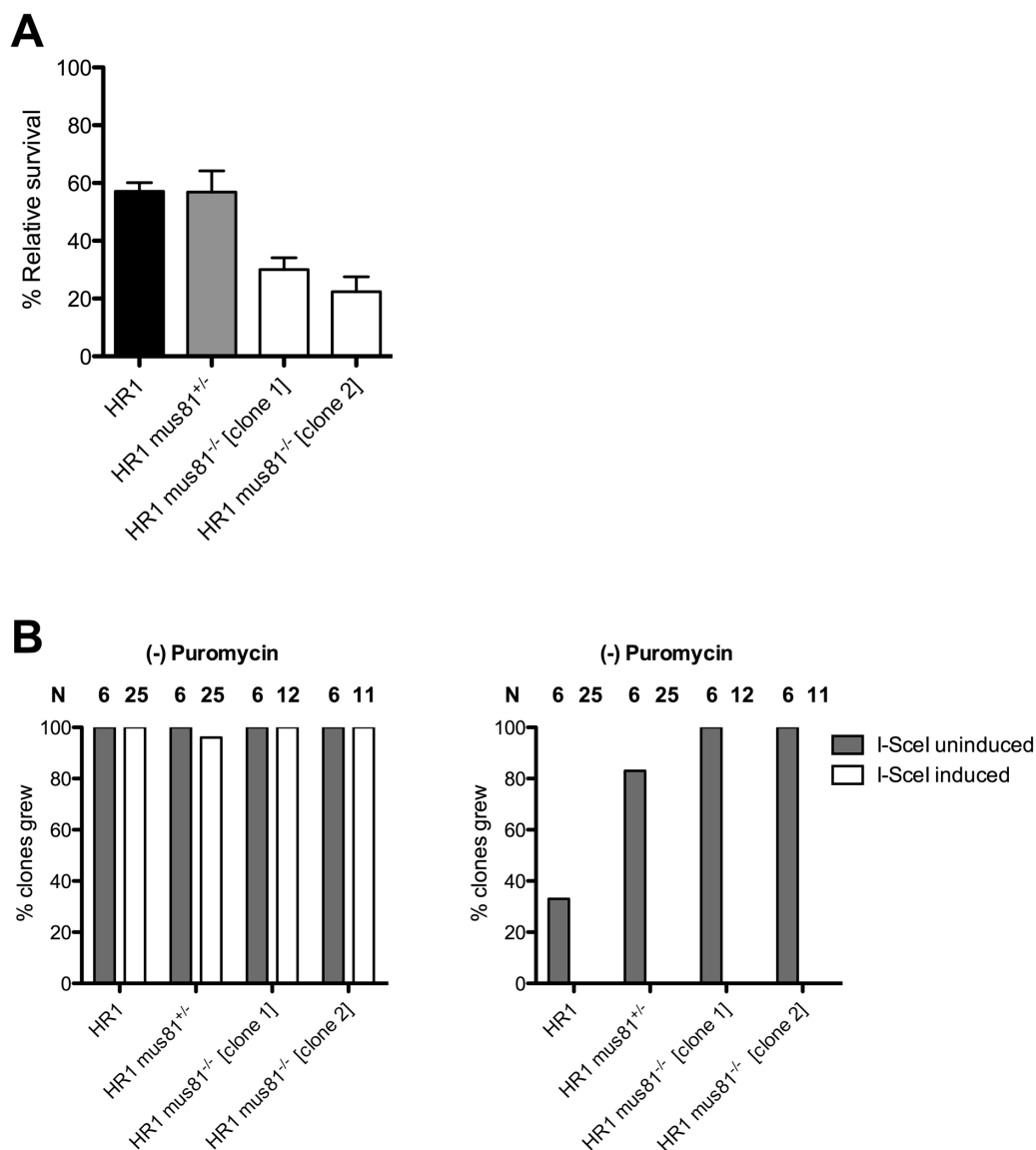
(A) Cell survival is shown as the average percentage of clonal survivors following I-SceI induction compared to survivors without I-SceI induction. Error bars represent standard error of the mean (SEM) between the three 96 well plates. (B) Puromycin sensitivity of surviving I-SceI induced and uninduced clones represented as the percentage of tested clones that grew in the presence and absence of 1  $\mu\text{g.mL}^{-1}$  puromycin. N, number of clones analysed.

#### 4.4.2 Analysis of chromosome-internal DSB repair in *mus81* mutants

The clonal survival of *mus81* mutants in an HR1 background was assayed using the same protocol outlined above (see Section 4.4). Figure 4-10a shows that the survival of HR1 wild type cells following DSB induction was 57.1%, similar to the survival rates obtained previously in this study (61.8%, see Section 4.4.1) and by Glover *et al.* (2008). Loss of one *MUS81* allele did not alter survival following DSB induction, with 56.9% of HR1 *mus81*<sup>+/-</sup> clones surviving. However, loss of both *MUS81* alleles resulted in an approximately two-fold decrease in survival

following I-SceI induction, with 30% and 22.4% survival rates seen in two different *HR1 mus81<sup>-/-</sup>* clones. The puromycin sensitivity status of surviving clones confirmed that most of the cell lines underwent DSB induction (Fig. 4-10B). A small number (8%) of *HR1 mus81<sup>-/-</sup> [clone 1]* induced survivors were puromycin resistant, suggesting they were not induced. However, the vast majority of *HR1 mus81<sup>-/-</sup> [clone 1]* survivors were puromycin sensitive as expected, and thus had undergone I-SceI cutting. Though all uninduced survivors should be puromycin resistant, only 33% of uninduced HR1 wild type survivors were resistant to puromycin (Fig. 4-9B). As discussed above (see Section 4.4.1), puromycin sensitivity in uninduced survivors could be due to leaky I-SceI expression leading to loss of the I-SceI recognition site and *PUR* in ‘uninduced’ cells. If correct, the smaller percentage (~20%) of uninduced *mus81<sup>+/-</sup>* clones that were puromycin sensitive may also have lost *PUR*.

These survival data of *HR1 mus81* mutants following DSB induction indicate that loss of *MUS81* impairs the repair of DSBs, though loss of one allele is insufficient to impair survival in this assay. The absence of a DSB repair phenotype following loss of one *MUS81* allele is consistent with clonal survival of *mus81* mutants following treatment with MMS and phleomycin. Furthermore, a role for *MUS81* in the repair of DSBs is consistent with the increased sensitivity of *mus81<sup>-/-</sup>* mutants to phleomycin and MMS (see Section 3.4.3.2).



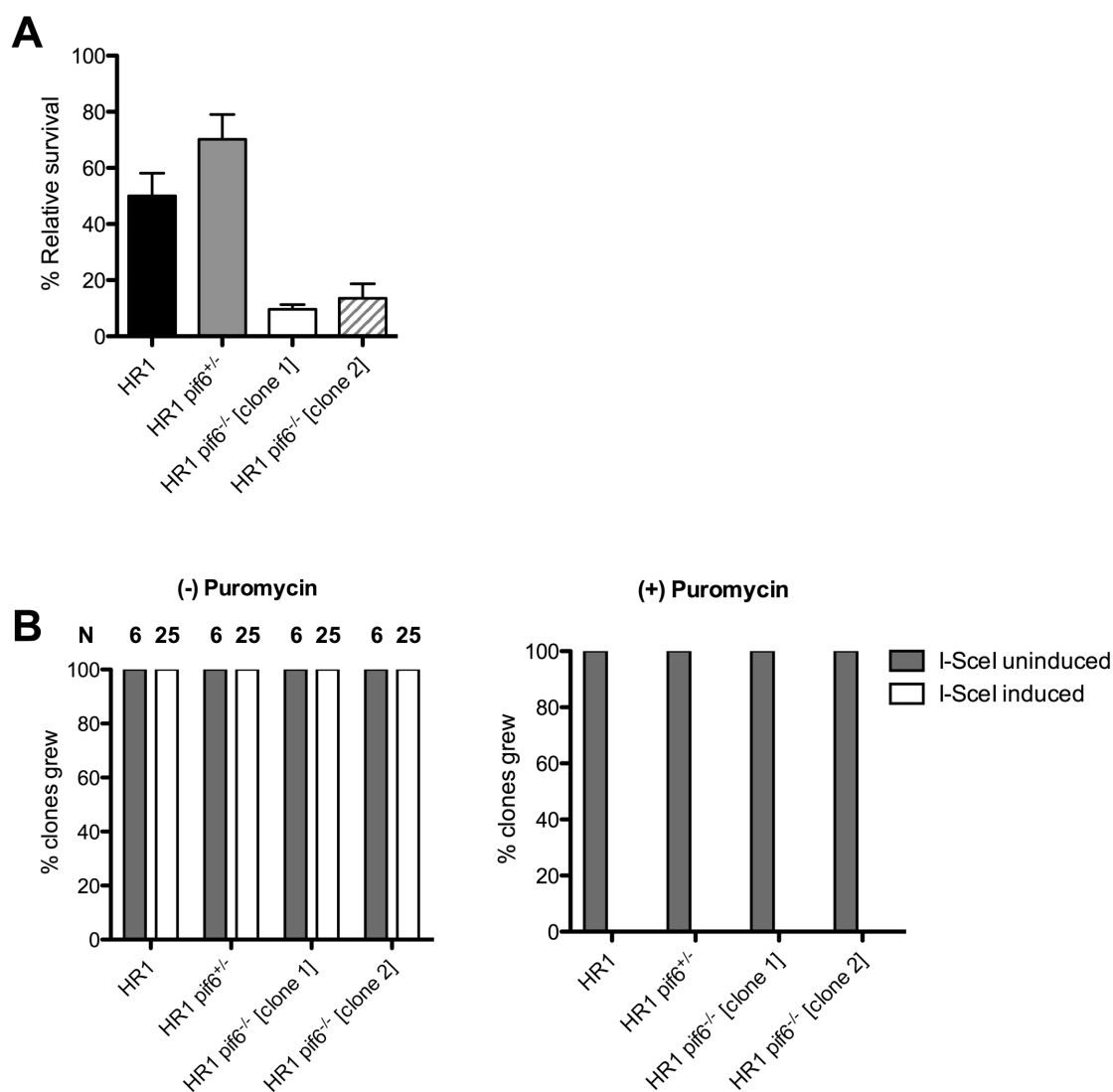
**Figure 4-9** Clonal survival following I-SceI induction in *HR1 mus81* mutants. *HR1* wild type, *HR1 mus81*<sup>+/-</sup> and two *HR1 mus81*<sup>-/-</sup> clones were distributed in three 96 well plates at a concentration of 0.26 cells per well either in the absence (I-SceI uninduced) or presence (I-SceI induced) of 2  $\mu\text{g.mL}^{-1}$  tetracycline. The number of wells with surviving cells (clones) was counted after 7-10 days of incubation. (A) Cell survival is shown as the average percentage of clonal survivors following I-SceI induction compared to survivors without I-SceI induction. Error bars represent standard error of the mean (SEM) between the three 96 well plates. (B) Puromycin sensitivity of surviving I-SceI induced and uninduced clones, represented as the percentage of tested clones that grew in the presence and absence of 1  $\mu\text{g.mL}^{-1}$  puromycin. N, number of clones analysed.

#### 4.4.3 Analysis of chromosome-internal DSB repair in *pif6* mutants

To investigate whether PIF6 is involved in DSB repair in *T. brucei*, the clonal survival of *pif6* mutants in the *HR1* background was assayed and clones analysed by PCR to attempt to determine the repair pathway profile. The clonal survival of *HR1* wild type cells was 50% (Fig. 4-10A), again similar to that observed previously in this work. Survival increased to 70% in the *HR1 pif6*<sup>+/-</sup> mutants but



decreased approximately five-fold in the two *HR1 pif6<sup>-/-</sup>* clones (9.6% and 13.5%). The puromycin sensitivity data for all survivors in this assay indicate that all ‘induced’ clones were puromycin sensitive and ‘uninduced’ clones were puromycin resistant as expected. Furthermore, all clones grew in the absence of puromycin. The puromycin sensitivity data therefore indicates that all ‘induced’ cultures were successfully induced (Fig. 4-10B).



**Figure 4-10 Clonal survival following I-SceI induction in *HR1 pif6* mutants**  
*HR1* wild type, *HR1 pif6*<sup>+/-</sup> and two *HR1 pif6*<sup>-/-</sup> clones were distributed in three 96 well plates at a concentration of 0.26 cells per well either in the absence (I-SceI uninduced) or presence (I-SceI induced) of 2 µg.mL<sup>-1</sup> tetracycline. The number of wells with surviving cells (clones) was counted after one week of incubation. (A) Cell survival is shown as the average percentage of clonal survivors following I-SceI induction compared to survivors without I-SceI induction. Error bars represent standard error of the mean (SEM) between the three 96 well plates. (B) Puromycin sensitivity of surviving I-SceI induced and uninduced clones, represented as the percentage of tested clones that grew in the presence and absence of 1 µg.mL<sup>-1</sup> puromycin. N, number of clones analysed.

The similarity in survival of HR1 wild type and *HR1 pif6<sup>+/-</sup>* mutants, compared with the reduced survival of the *pif6<sup>-/-</sup>* mutants, following a DSB is distinct from the survival profiles of these mutants following more general DNA damage. In previous survival assays of *pif6* mutants after treatment with the DNA damaging agents hydroxyurea (HU) and MMS, *pif6<sup>+/-</sup>* and *pif6<sup>-/-</sup>* displayed very similar or indistinguishable survival phenotypes (see Section 3.4.3.2). Both mutants were more resistant to MMS, an inducer of DSBs, and less resistant to HU, primarily a replication-stalling agent. It might therefore be predicted that both *HR1 pif6<sup>+/-</sup>* and *HR1 pif6<sup>-/-</sup>* mutants would display an increased resistance to DSB induction. In contrast, *HR1 pif6<sup>+/-</sup>* mutants were comparable with wild type and *HR1 pif6<sup>-/-</sup>* mutants displayed impaired survival, a different phenotype from the other cells.

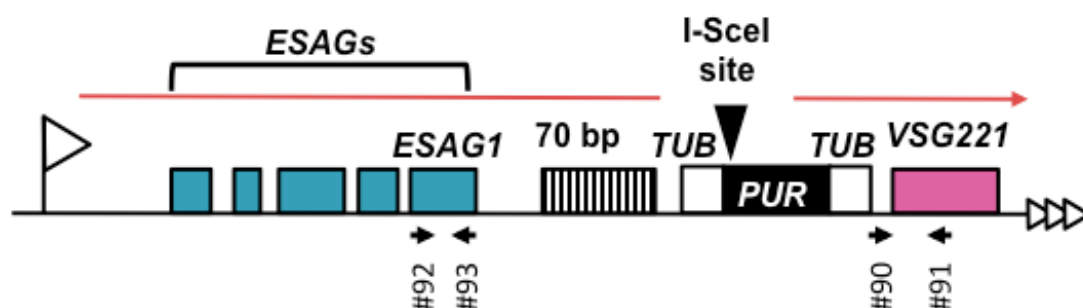
These data suggest that although the loss of one or both *PIF6* alleles is sufficient to increase survival following MMS treatment, loss of a single *PIF6* allele does not alter survival following a DSB at a chromosomal internal location. Furthermore, whereas *pif6<sup>-/-</sup>* mutants are more resistant to ‘generalised’ DNA damage induced by MMS or hydroxyurea, the same null mutants are considerably impaired in their ability to repair an induced DSB at a chromosomal internal location. Thus, the role of PIF6 in genome maintenance or repair may be quite different in response to different lesions.

## 4.5 Analysis of DSB repair in the active VSG expression site

The clonal survival of *recq2*, *mus81* and *pif6* mutants in an HRES background was next analysed in order to investigate whether these factors influence repair of DSBs upstream of the active VSG, the region in which DNA breaks initiating VSG switching are hypothesised to occur. If mutants display an altered ability to survive a DSB in this location this can then be correlated with whether or not the same mutation also affects VSG switching, therefore testing whether DSBs play a role in VSG switching. Such a correlation has been demonstrated for RAD51, mutants of which have a lower rate of VSG switching (McCulloch & Barry, 1999). Survival of *HRES rad51<sup>-/-</sup>* cells following I-SceI induction is also impaired compared to wild type HRES cells (Glover *et al.*, 2013a). In addition, this assay allows us to compare whether or not the same strategies for DSB repair are used in response to a DSB at the site hypothesised to initiate VSG switching relative to

a chromosome-internal DSB (in the HR1 background). To date, different responses in these two locations have only been observed in mutants of a *T. brucei* histone acetyltransferase (HAT3) (Glover & Horn, 2014).

In HRES cell lines a DSB adjacent to the 70 bp repeats has been demonstrated to induce VSG switching (Boothroyd *et al.*, 2009; Glover *et al.*, 2013a). Surviving clones were therefore assayed by PCR for presence of VSG221 using primers #90 and #91 (955 bp PCR product). As VSG221 is lost in almost all switched survivors following I-SceI induction in the HRES line (Glover *et al.*, 2013a), this PCR can therefore be used to confirm that surviving clones have switched. Surviving clones were also assayed by PCR for presence of the ESAG1 present in and unique to the VSG221 active VSG BES using primers #92 and #93 (328 bp product) (Fig. 4-11). ESAG1 is located upstream of the 70 bp repeats and its presence or absence in switched clones can be used to help determine the kind of switching event. In most switched survivors ESAG1 is retained (Glover *et al.*, 2013a), indicating that the VSG BES sequences upstream of the I-SceI site are still present and that recombination initiated in the region of the 70 bp repeats. ESAG1 absence indicates that recombination initiated upstream of the 70 bp repeats or switched through loss or replacement of the VSG BES. The results of all PCR analyses on assays described below can be found in Appendix 7.1.



**Figure 4-11** Primer binding sites for PCR analysis of HRES I-SceI assay clones. The VSG221 active BES is shown. To test for presence of VSG221, a 328 bp PCR product was generated, using primers #90 and #91, which bind upstream of and within the VSG221 ORF. To test for presence of ESAG1, a 955 bp region of the ESAG1 ORF was PCR amplified, using primers #92 and #93. Promoter (flag), red line indicates expression of the VSG BES, ESAGs (turquoise boxes), 70 bp repeats (hatched box), VSG221 (pink box), triangles (telomeric repeats). Primer sequences can be found in Appendix 7.1. Not to scale.

### 4.5.1 Analysis of active VSG expression site DSB repair in *recq2* mutants

It has previously been reported that survival of HRES wild type cells after I-SceI induction was extremely low at around 5% of uninduced cells and 10-fold lower than HR1 induced cells (Glover *et al.*, 2013a). Figure 4-12A shows that the survival of HRES wild type cells (23.8%) in this assay was two-fold lower than that of HR1 cells (see Section 4.4) and around four to five-fold higher than reported by Glover *et al.* (2013a). This difference in survival could be due to differences in culturing conditions, such as the culture media used here containing a higher concentration of FBS (20% vs 10%). Regardless, the survival rate of the two *HRES recq2*<sup>-/-</sup> clones (13% and 14.8%) was approximately half that of HRES wild type cells, suggesting absence of RECQ2 impaired survival (Fig. 4-12A).

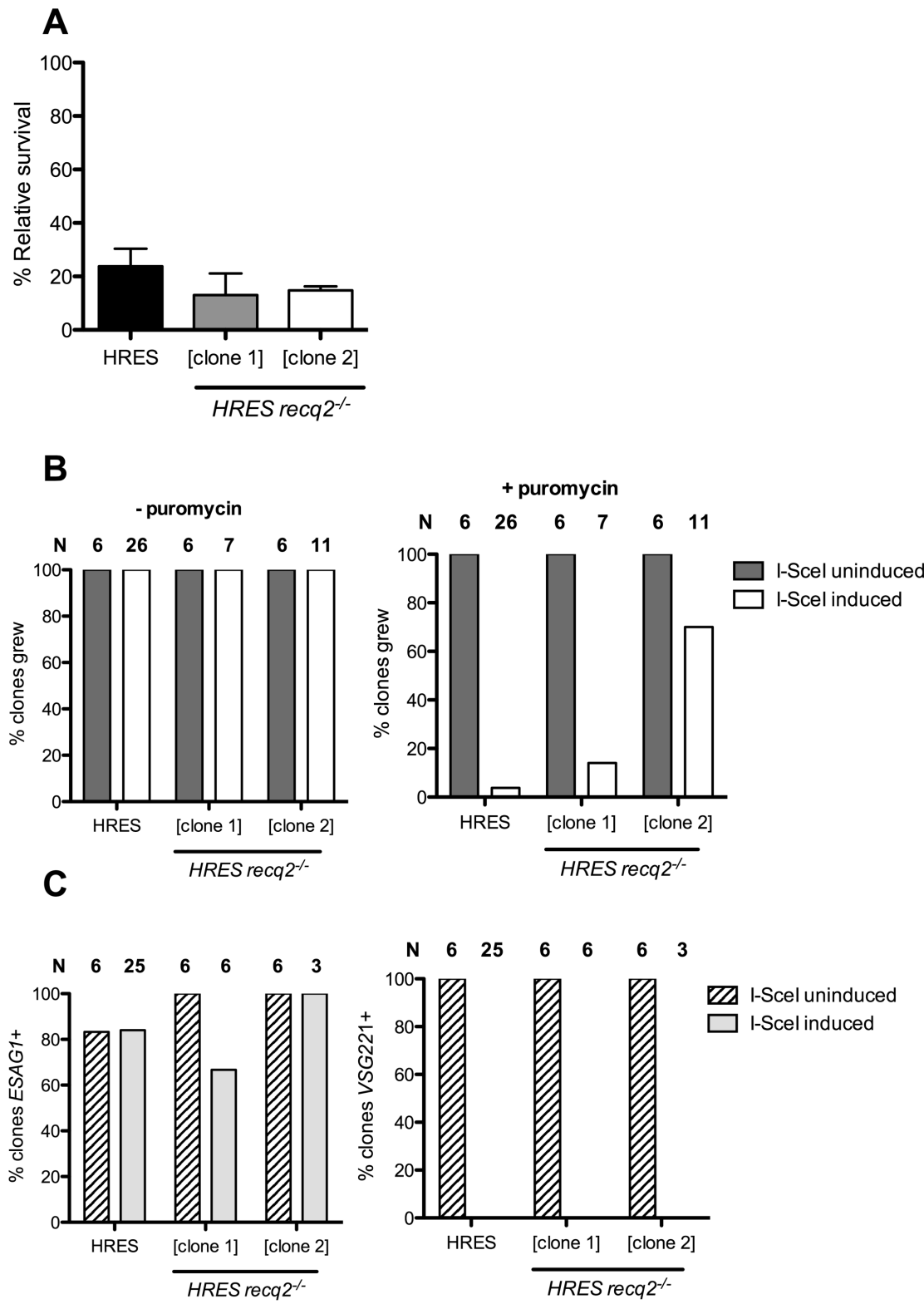
The puromycin resistance data (Fig. 4-12B) indicate that not all cells were successfully induced. A small proportion (3.8%) of HRES wild type clones recovered were puromycin resistant. A higher frequency of resistance was seen in the mutants: 14% (one out of seven clones analysed) of *HRES recq2*<sup>-/-</sup> [clone 1] and 70% (seven out of ten clones analysed) of *HRES recq2*<sup>-/-</sup> [clone 1] survivors were puromycin resistant. The increased incidence of puromycin-resistant cells in the mutants may reflect the impaired survival, meaning that a larger proportion of cells in the population not subjected to DSB induction were recovered on cloning. These clones were excluded from the *VSG221* and *ESAG1* PCR analysis of survivors shown in Figure 4-12C. All other surviving induced clones lost *VSG221* as expected and *ESAG1* was retained in over 80% of HRES wild type survivors (25 clones analysed), consistent with Glover *et al.*, (2013a). *ESAG1* loss was higher in *HRES recq2*<sup>-/-</sup> [clone 1] survivors (~40%) but the total number of clones analysed was much lower (six). *ESAG1* was retained by all survivors of the second *HRES recq2*<sup>-/-</sup> clone analysed. It is therefore not possible to draw firm conclusions regarding whether *ESAG1* loss is greater in *HRES recq2*<sup>-/-</sup> mutants. However, it seems likely that the profile of *VSG221* loss and *ESAG1* retention is seen in both wild type and *HRES recq2*<sup>-/-</sup> mutants after I-SceI induction, despite the differing survival. Details of the diagnostic PCRs and puromycin testing for each clone analysed can be found in Appendix 7.4.

The lowered survival of HRES *recq2*<sup>-/-</sup> mutants after induction of a DSB, relative to wild type HRES cells, suggests that RECQ2 is involved in the repair of DSBs in the active VSG BES. However, the survival of *recq2*<sup>-/-</sup> mutants was not disproportionately lower in HRES than HR1 after I-SceI induction (a survival impairment of ~50% relative to wild type was seen in both cell lines), indicating that the role of RECQ2 in DSB repair is not specific to the active VSG BES. A role for RECQ2 in the repair of DSBs at the active VSG BES is compatible with clonal survival assays (see Section 3.4.3.2), which showed that *recq2* mutants are more sensitive to DNA damaging agents that cause DSBs, MMS and phleomycin. In addition, it is compatible with the observation that myc-tagged RECQ2 relocalises to detectable subnuclear foci after phleomycin treatment.

As for the HR1 experiment, the survival of *HRES recq2* mutants was assayed a second time (Fig. 4-13A), again with differing results to the first experiment. The survival rate of HRES wild type cells (13.1%) was lower than the first time the experiment was carried out, though similar to the survival rate in other repeats of this assay, described below (see Sections 4.5.2 and 4.5.3). In addition to the HRES wild type and *HRES recq2*<sup>-/-</sup> cell lines assayed previously, the survival rate of *HRES recq2*<sup>+/-</sup> was also examined. In contrast to the previous experiment, survival following I-SceI induction was not altered in *HRES recq2*<sup>-/-</sup> cells and furthermore, the survival rate of *HRES recq2*<sup>+/-</sup> cells was 24%, almost twice that of wild type HRES.

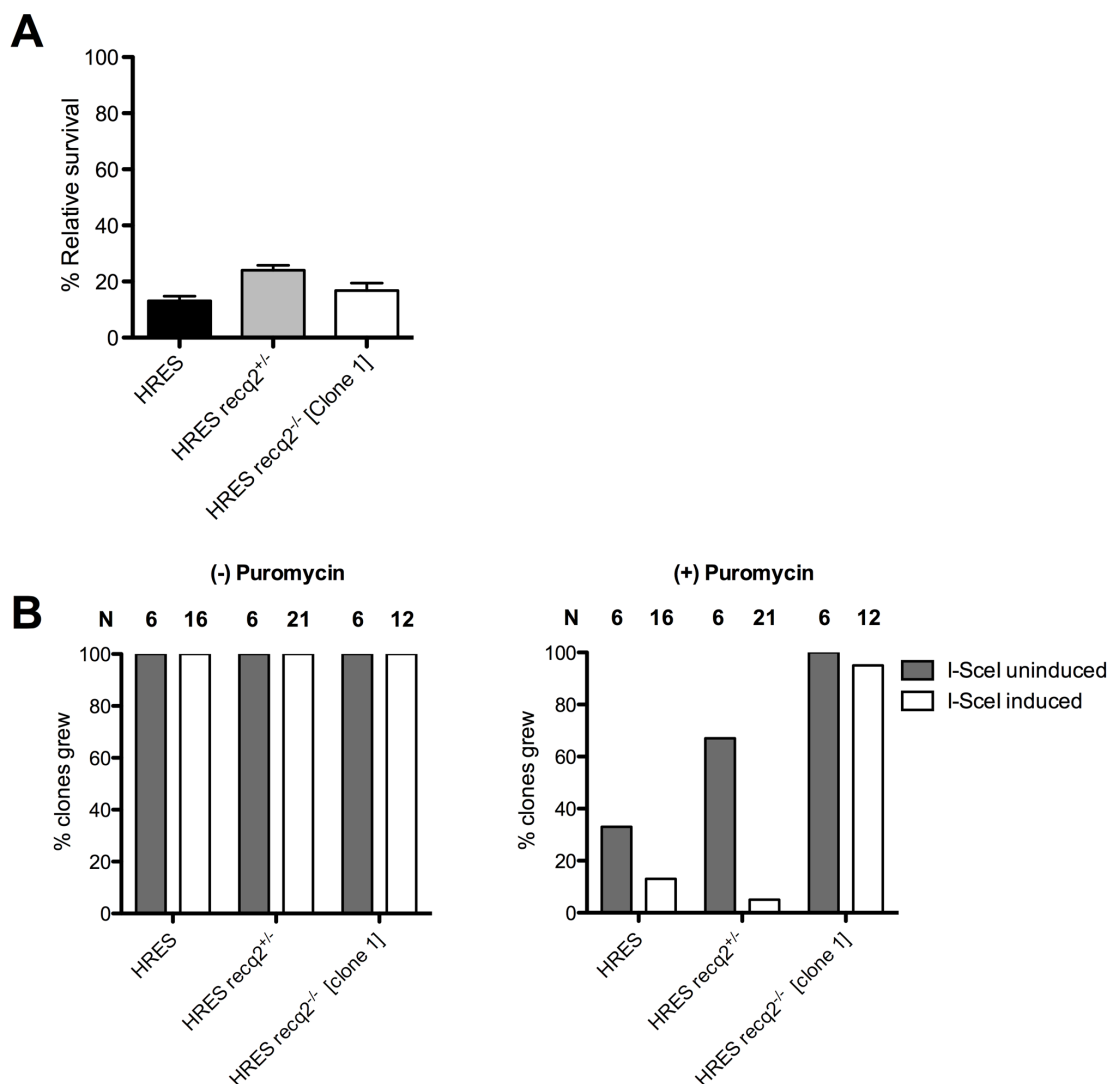
These data are inconsistent with the impaired survival of *HRES recq2* mutant survival in the first assay, impaired survival of *HR1 recq2* mutants, and the sensitivity of *recq2* mutants to a range of DNA damaging agents. Furthermore, as seen with HR1, assessing the puromycin resistance data for the surviving clones of this experiment suggest the same difficulties as described for the HR1 cells. Figure 4-11B shows that 67% of uninduced HRES wild type and 33% of *HRES recq2*<sup>+/-</sup> survivors were puromycin sensitive, when previously no such cells were seen. Additionally, puromycin resistant clones were seen for the induced HRES wild type and *HRES recq2*<sup>+/-</sup> cells, and nearly all of the *HRES recq2*<sup>-/-</sup> survivors were puromycin resistant, indicating that none of these cultures had a 100% I-SceI induction rate. As before, this experiment was conducted after the first experiment, and after cells had been in culture for a longer period of time, providing further evidence that continued growth of either HR1 or HRES cells

leads to instability. The problems with determining whether cultures were induced or not in this experiment, and lack of consistency with previous data on *RECQ2*, mean that this second assay cannot be relied upon.



**Figure 4-12** Clonal survival and *ESAG1* and *VSG221* loss following I-SceI induction in *HRES recq2* mutants

**HRES wild type and two *HRES recq2*<sup>-/-</sup> clones were distributed in three 96 well plates at a concentration of 0.56 cells per well either in the absence (I-SceI uninduced) or presence (I-SceI induced) of 2  $\mu\text{g.mL}^{-1}$  tetracycline. The number of wells with surviving cells (clones) was counted after one week of incubation. (A) Cell survival is shown as the average percentage of clonal survivors following I-SceI induction compared to survivors without I-SceI induction. Error bars represent standard error of the mean (SEM) between the three 96 well plates. (B) Puromycin sensitivity of surviving I-SceI induced and uninduced clones, represented as the percentage of tested clones that grew in the presence and absence of 1  $\mu\text{g.mL}^{-1}$  puromycin. (C) Surviving I-SceI uninduced and induced clones from (B), excluding I-SceI induced clones that were puromycin resistant, were assayed for *ESAG1* and *VSG221* presence by PCR. N, number of clones analysed.**



**Figure 4-13 Repeat of analysis of clonal survival following I-SceI induction in *HRES recq2* mutants**

**HRES wild type, *HRES recq2*<sup>+/+</sup> and *HRES recq2*<sup>-/-</sup> cells were distributed in three 96 well plates at a concentration of 0.56 cells per well either in the absence (I-SceI uninduced) or presence (I-SceI induced) of 2  $\mu\text{g.mL}^{-1}$  tetracycline. The number of wells with surviving cells (clones) was counted after 7-10 days incubation. (A) Cell survival is shown as the average percentage of clonal survivors following I-SceI induction compared to survivors without I-SceI induction. Error bars represent standard error of the mean (SEM) between the three 96 well plates. (B) Puromycin sensitivity of surviving I-SceI induced and uninduced clones, represented as the percentage of tested clones that grew in the presence and absence of 1  $\mu\text{g.mL}^{-1}$  puromycin. N, number of clones analysed.**

### 4.5.2 Analysis of active VSG expression site DSB repair in *mus81* mutants

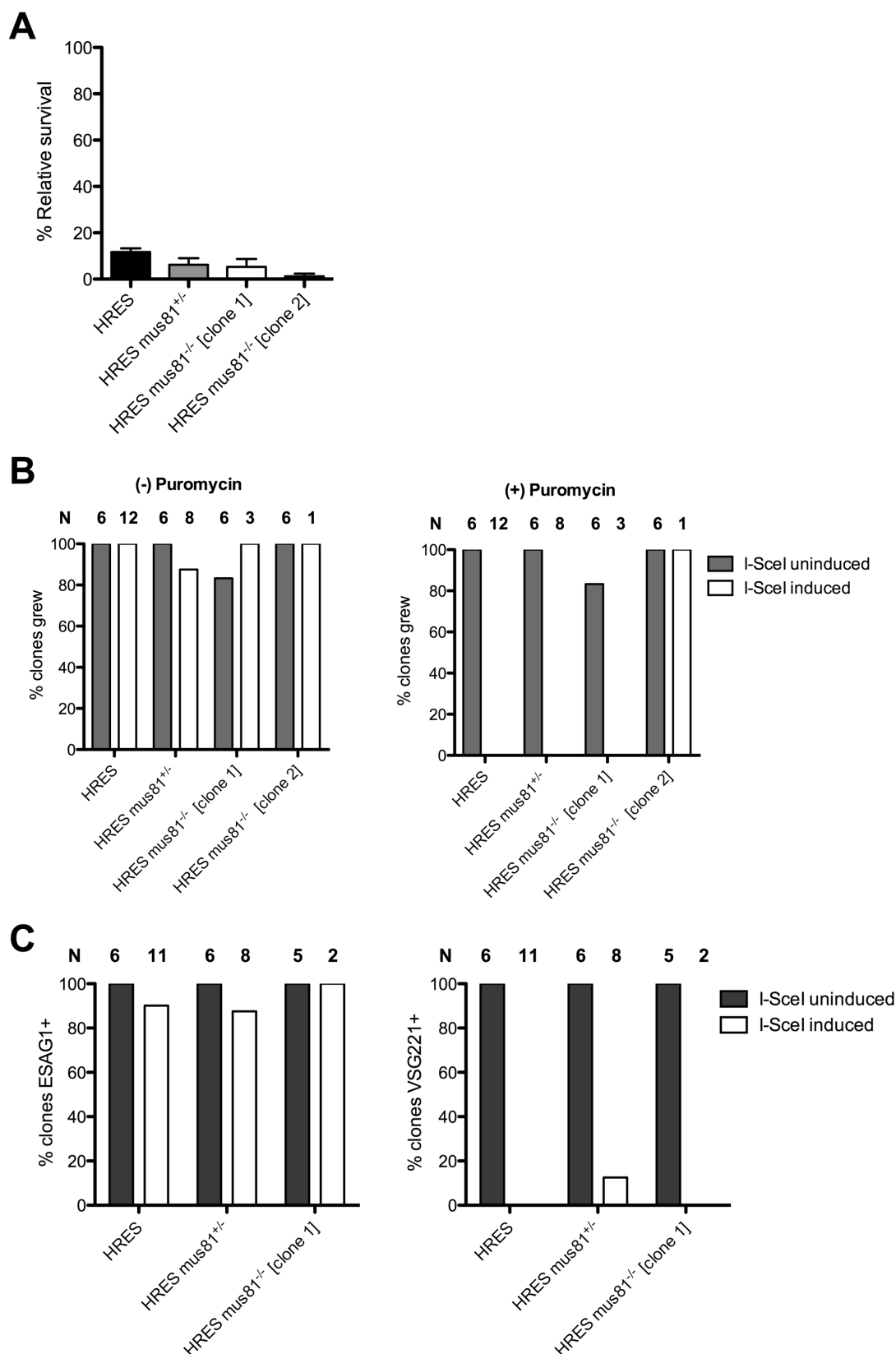
The ability of *mus81* mutants to survive the induction of a DSB in the active VSG BES was next analysed, assaying an *HRES mus81<sup>+/-</sup>* mutant and two *HRES mus81<sup>-/-</sup>* mutants. *HRES mus81<sup>+/-</sup>* mutants displayed a survival rate (30%) two-fold lower than HRES wild type cells (60%) (Fig. 4-14A). This contrasts with the phenotype of *HR1 mus81<sup>+/-</sup>* mutants, which exhibited an unchanged survival rate from HR1 wild type following I-SceI induction. It is also different to the phenotype observed in *mus81<sup>+/-</sup>* mutants following treatment with phleomycin and MMS, both of which cause DSBs as well as other types of DNA damage. *HRES mus81<sup>-/-</sup>* mutants had impaired survival following I-SceI-induced DSBs. This was observed in both *mus81<sup>-/-</sup>* mutants analysed, with survival rates of 5.3% and 1.2% following I-SceI induction, respectively (Fig. 4-14A). The puromycin sensitivity of surviving clones confirms that all ‘induced’ HRES wild type, *HRES mus81<sup>+/-</sup>* and *HRES mus81<sup>-/-</sup> [clone 1]* survivors underwent I-SceI induction. Only a single induced *HRES mus81<sup>-/-</sup> [clone 2]* survivor was recovered in this experiment, and this was puromycin resistant, indicating that a DSB may not have been induced in that cell.

PCR analysis of survivors revealed that all induced *HRES* wild type and *HRES mus81<sup>-/-</sup> [clone 1]* survivors lost *VSG221*, while a minority (12.5%) of *HRES mus81<sup>+/-</sup>* survivors retained *VSG221* (Fig. 4-14B). As all induced survivors were puromycin sensitive and therefore underwent I-SceI induction and DSB formation, it is possible that these *VSG221* positive survivors underwent repair of the induced DSB through telomere exchange, meaning the *VSG221* gene is located in another VSG BES. However, this most likely reflects sampling variation rather than mutation of *MUS81*, as the same pattern was not seen in the *mus81<sup>-/-</sup>* mutants. *ESAG1* was retained by >80% of all induced survivors (Fig. 4-14B), with no clear differences between wild type and mutants, indicating that recombination initiated downstream of *ESAG1* in virtually all induced survivors. Genomic DNA was not recovered for the single induced *HRES mus81<sup>-/-</sup>* survivor in this assay and thus PCR analysis of *ESAG1* and *VSG221* presence was not possible.

Decreased ability of *mus81<sup>-/-</sup>* cells to survive a DSB induced adjacent to the 70 bp repeats suggests that *MUS81* is involved in the repair of DSBs in this



location, in addition to playing a role in the repair of DSBs elsewhere in the chromosome and in repair of phleomycin and MMS-induced DNA damage.



**Figure 4-14** Clonal survival and *ESAG1* and *VSG221* loss following I-SceI induction in *HRES mus81* mutants

*HRES* wild type, *HRES mus81*<sup>+/-</sup> and two *HRES mus81*<sup>-/-</sup> clones were distributed in three 96 well plates at a concentration of 0.56 cells per well either in the absence (I-SceI uninduced) or presence (I-SceI induced) of 2  $\mu\text{g.mL}^{-1}$  tetracycline. The number of wells with surviving cells (clones) was counted after one week of incubation. (A) Cell survival is shown as the average percentage of clonal survivors following I-SceI induction compared to survivors

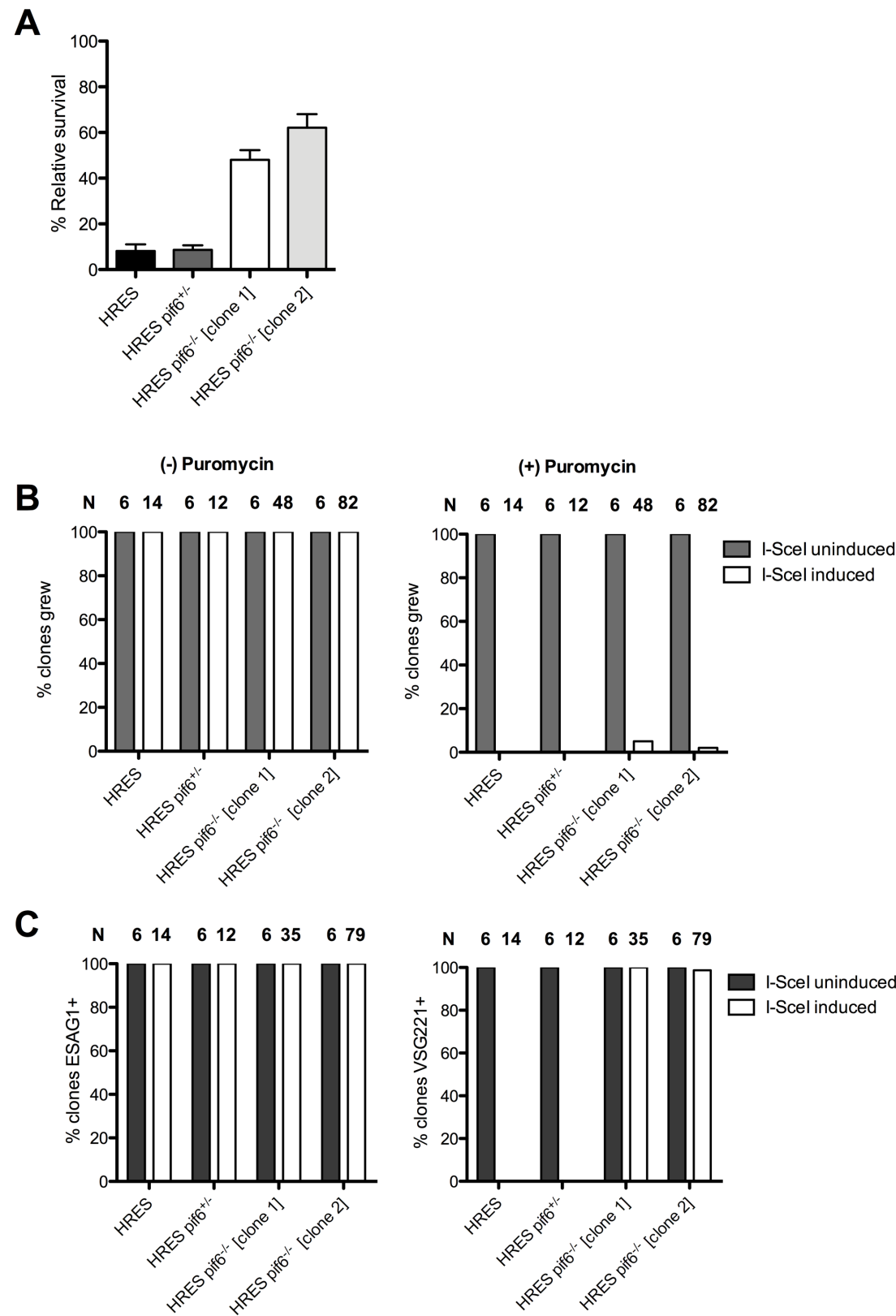
without I-SceI induction. Error bars represent standard error of the mean (SEM) between the three 96 well plates. (B) Puromycin sensitivity of surviving I-SceI induced and uninduced clones, represented as the percentage of tested clones that grew in the presence and absence of  $1 \mu\text{g.mL}^{-1}$  puromycin. (C) Recovered clones from (B) were assayed for *ESAG1* and *VSG221* presence by PCR. Genomic DNA was not recovered from one induced *HRES* wild type clone, one uninduced *HRES mus81<sup>-/-</sup>* [clone 1] and one induced *HRES mus81<sup>-/-</sup>* [clone 1] clone - these are excluded from the PCR analysis shown in (C). The single induced *HRES mus81<sup>-/-</sup>* [clone 2] recovered was puromycin resistant and therefore *HRES mus81<sup>-/-</sup>* [clone 2] survivors were excluded from PCR analysis. N, number of clones analysed.

### 4.5.3 Analysis of active expression site DSB repair in *pif6* mutants

In contrast with the increased sensitivity of *mus81* and *recq2* mutants to induced DSBs in HRES, analysis of the survival of *HRES pif6* mutants following I-SceI induction revealed a distinct phenotype. The survival of induced HRES wild type cells was comparable with previous assays (8.1%) and loss of one *PIF6* allele in the *HRES pif6<sup>+/-</sup>* cells did not alter the survival rate (8.6%) (Fig. 4-15A). However, there was a striking increase in survival in both *HRES pif6<sup>-/-</sup>* clones analysed. Compared to HRES wild type, *HRES pif6<sup>-/-</sup>* [clone 1] survival increased six-fold (to 48%), while survival of the second clone increased over seven-fold (to 62.1%), (Fig. 4-15A). The puromycin resistance data (Fig. 4-16B) showed that all HRES wild type and *HRES pif6<sup>+/-</sup>* 'induced' clones were puromycin sensitive and thus underwent I-SceI cutting. The high *HRES pif6<sup>-/-</sup>* survival was not due to inefficient I-SceI induction, as >95% of surviving *HRES pif6<sup>-/-</sup>* induced survivors were puromycin sensitive and thus were successfully induced. This suggests that the increase in *HRES pif6<sup>-/-</sup>* survival was not a result of unforeseen alterations in the I-SceI target or gene. To attempt to test this further, given the striking difference in this phenotype relative to the other mutants and to *pif6<sup>-/-</sup>* HR1 mutants, PCR-amplification of the *PUR* gene was also attempted in order to investigate if it was removed by the DSB induction. However, multiple attempts, including in the uninduced controls, produced inconsistent results, likely due to the high GC content of the *PUR* gene and therefore these data are not presented here. Nonetheless, the increase in *HRES pif6<sup>-/-</sup>* survival is comparable with *pif6* mutant survival following MMS damage, though differs in that *pif6<sup>+/-</sup>* mutants also displayed increased survival after MMS treatment. A phenotypic separation between the two mutants was also seen in the I-SceI assays in HR1 cells, where *pif6<sup>+/-</sup>* mutants behaved more like wild type cells and were again different from *pif6<sup>-/-</sup>*.

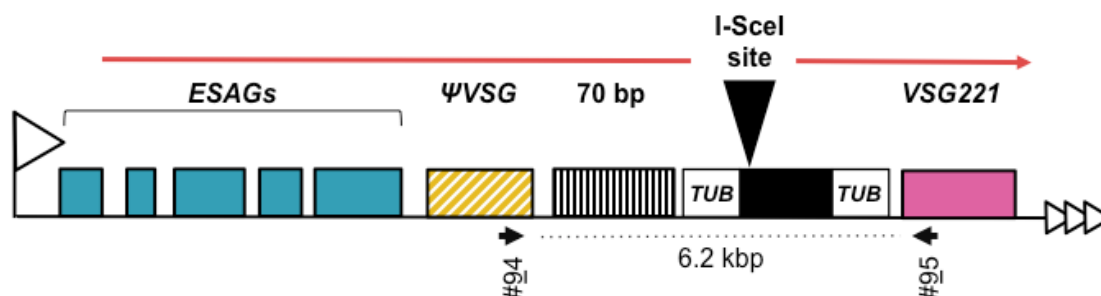
Analysis of *VSG221* and *ESAG1* presence by PCR revealed a further unexpected result (Fig. 4-15C). DSB induction near the 70 bp repeats causes loss of *VSG221* in almost all HRES survivors, as observed in assays described above (including for *recq2* and *mus81* mutants) and elsewhere (Glover *et al.*, 2013a). Though *VSG221* was lost from all induced HRES wild type and *HRES pif6<sup>+/-</sup>* survivors as predicted, it was retained in >98% of all induced *HRES pif6<sup>-/-</sup>* survivors. Conversely, *ESAG1* was retained in all induced survivors. Amplification of *VSG221* by PCR showed that the gene is present but does not reveal its location. For instance, induced survivors in which *VSG221* was retained could have repaired the break and switched VSG via telomere exchange, thus the *VSG221* gene would be in a different genomic location. Alternatively, *VSG221* may be retained in the VSG BES and may, indeed, continue to be expressed, meaning that in these cells induction of a DSB might not elicit a VSG switch. It was therefore attempted to determine the genomic location of *VSG221* in all survivors using PCR. A pair of PCR primers was designed that amplified a region of the active VSG BES, the forward primer binding within a VSG pseudogene (accession number H25N7.32, tritrypdb.org) upstream of the 70 bp repeats, unique to the active VSG BES, and a reverse primer binding in *VSG221*. Successful amplification of this PCR product (6.2 kbp in uninduced clones) would most likely indicate that *VSG221* is located in the active VSG BES, whereas a negative result would suggest that *VSG221* is in a different genomic location. Multiple attempts at this PCR using a range of conditions and polymerases were unsuccessful and are not presented here. This is most likely due to the length of the PCR product and the presence of the *PUR* gene.

The clonal survival assay of *HRES pif6* mutants was repeated an additional two times, selecting one of the *HRES pif6<sup>-/-</sup>* clones for analysis. The results of these assays were very similar and are shown combined in Figure 4-17A. Similar to the assay described above, following induction, the survival rate of HRES wild type cells was 15.6% and the survival of *HRES pif6<sup>+/-</sup>* cells was unchanged compared to wild type (15%). However, the increase in induced *HRES pif6<sup>-/-</sup>* survival observed in the first experiment was not replicated. The survival rate of induced *HRES pif6<sup>-/-</sup>* cells was similar to that of *HRES* wild type and *HRES pif6<sup>+/-</sup>* (18.5%).



**Figure 4-15 Analysis of clonal survival and *ESAG1* and *VSG221* loss following I-SceI induction in *HRES pif6* mutants**  
*HRES* wild type, *HRES pif6*<sup>+/-</sup> and two *HRES pif6*<sup>-/-</sup> clones were distributed in three 96 well plates at a concentration of 0.56 cells per well either in the absence (I-SceI uninduced) or presence (I-SceI induced) of 2 µg.mL<sup>-1</sup> tetracycline. The number of wells with surviving cells (clones) was counted after one week of incubation. (A) Cell survival is shown as the average percentage of clonal survivors following I-SceI induction compared to survivors without I-SceI induction. Error bars represent standard error of the mean (SEM) between

the three 96 well plates. (B) Puromycin sensitivity of surviving I-SceI induced and uninduced clones, represented as the percentage of tested clones that grew in the presence and absence of  $1 \mu\text{g.mL}^{-1}$  puromycin. (C) Clones from (B), excluding induced clones that were puromycin resistant or clones that were assayed for *ESAG1* and *VSG221* presence by PCR. N, number of clones analysed.

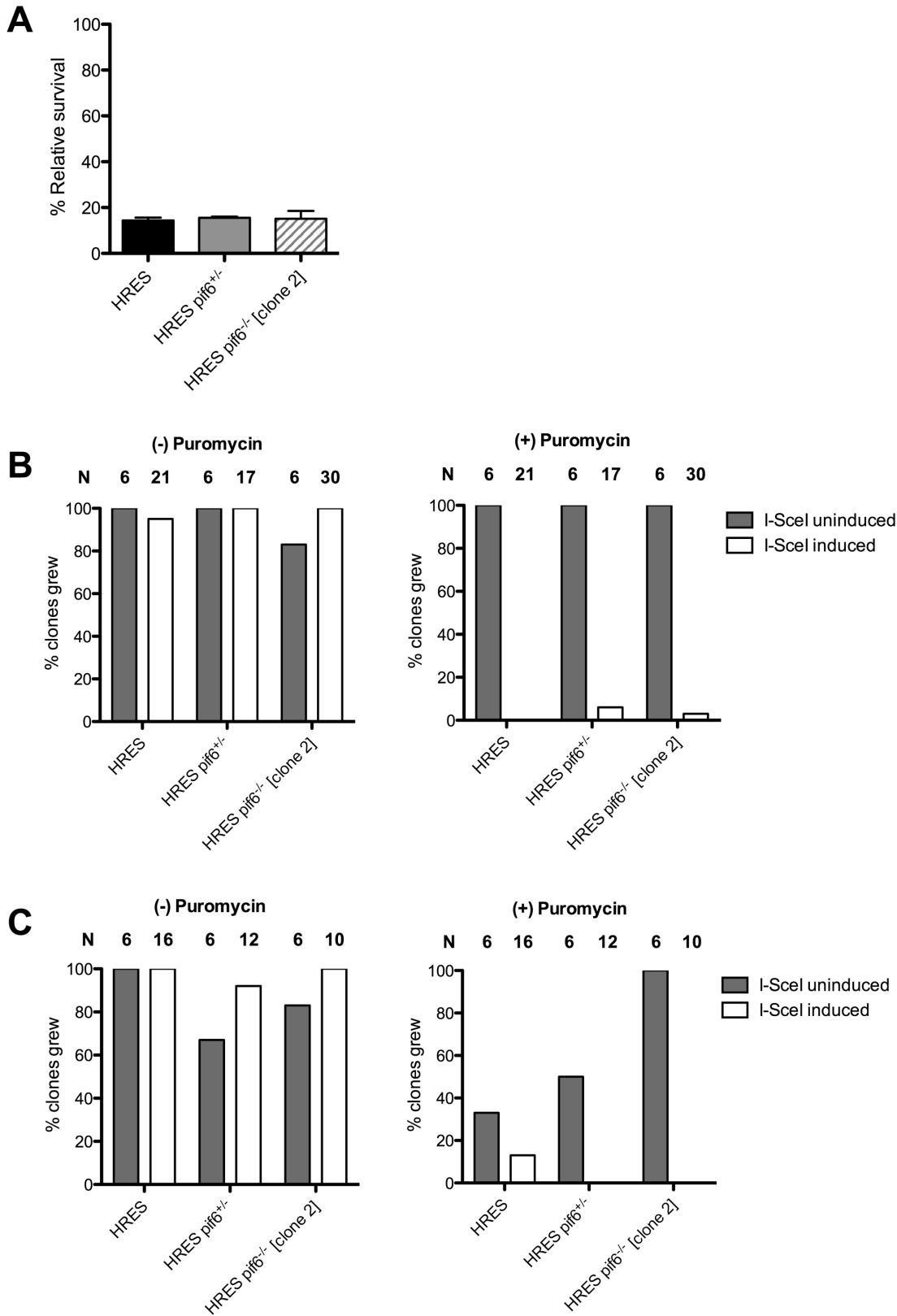


**Figure 4-16 Attempt by PCR to analyse VSG221 genomic location**  
A schematic of the VSG221 active VSG BES is shown. Black arrows indicate primer binding locations. PCR amplification to determine the genomic location of VSG221 was attempted using a primer (primer #94) that bound in the ORF of the VSG pseudogene ( $\Psi$ VSG, yellow hatched box; H25N7.34, trytripdb.org) unique to the VSG221 BES and a primer that bound in the VSG221 ORF (primer #95), predicted product 6.2 kbp. Sequences for primers can be found in Appendix 7.1. Not to scale.

However, the above combined dataset suffers from problems with the puromycin sensitivity data that have been described above and appear to be manifest when using more established cells. In one experiment in this dataset, all uninduced survivors displayed puromycin resistance and only a small proportion ( $\leq 6\%$ ) of induced survivors were puromycin resistant, indicating that most cells in ‘induced’ cultures were successfully induced (Fig. 4-17b). However, the second experiment in this dataset (dataset 2) suffered from multiple problems in the puromycin resistance profile. Some uninduced clones were not resistant to puromycin and some induced clones were resistant to puromycin (Fig. 4-17c). Only 33% of HRES *pif6*<sup>+/-</sup> uninduced clones and 50% of HRES *pif6*<sup>-/-</sup> uninduced clones were puromycin resistant. In addition, 13% of HRES wild type induced clones were resistant to puromycin. Even in the absence of puromycin, 33% of HRES *pif6*<sup>+/-</sup> uninduced survivors and 17% of HRES *pif6*<sup>-/-</sup> uninduced survivors did not grow.

Due to the reasons outlined above, the combined data in the experimental repeats (Fig. 4-17) are less reliable than the first experiment (Fig. 4-15), almost certainly for the same reasons of I-SceI-induced instability seen in all such repeats. Nonetheless, in considering the second combined dataset (Fig. 4-17), and although the increased HRES *pif6*<sup>-/-</sup> survival observed in the first experiment

was not replicated, loss of PIF6 did not result in impaired survival following I-SceI induction in HRES as was seen in *recq2* and *mus81* mutants.



**Figure 4-17** Combined data of two repeats of clonal survival following I-SceI induction in *HRES pif6* mutants

HRES wild type, *HRES pif6*<sup>+/−</sup> and one of the *HRES pif6*<sup>−/−</sup> clones used in the previous assay (Fig. 4-15) were distributed in three 96 well plates at a concentration of 0.56 cells per well either in the absence (I-SceI uninduced) or presence (I-SceI induced) of 2  $\mu\text{g.mL}^{-1}$  tetracycline. The number of wells with surviving cells (clones) was counted after one week of incubation. (A) Cell survival is shown as the average percentage of clonal survivors following I-SceI induction compared to survivors without I-SceI induction. Error bars represent standard error of the mean (SEM) of the two datasets. (B & C) Puromycin sensitivity of surviving I-SceI induced and uninduced clones from dataset 1 (B) and dataset 2 (C), represented as the percentage of tested clones that grew in the presence and absence of 1  $\mu\text{g.mL}^{-1}$  puromycin. N, number of clones analysed.

The increase in *pif6*<sup>−/−</sup> survival following I-SceI induction in HRES in the first experiment (or the lack of impaired survival in the second experiment) is compatible with the observation that loss of PIF6 increases survival after MMS and phleomycin treatment. If this correlation is correct, then it is unclear why increased survival was not observed in *HR1 pif6*<sup>−/−</sup> mutants and, instead, the opposite effect was observed. Some PIF1 proteins in other eukaryotes have functions in telomere repair. For example, *S. cerevisiae* PIF1 (ScPif1) inhibits the addition of telomeric repeats to DSBs and, in its absence, the rate of telomere addition to DSBs increases (Makovets & Blackburn, 2009; Schulz & Zakian, 1994). Telomere addition to a telomere-proximal DSB (as in HRES) would result in loss of far fewer genes than telomere addition to a chromosome internal DSB (as in HR1). Such a function for PIF6 could then explain the different survival phenotypes of *pif6*<sup>−/−</sup> mutants in HR1 and HRES. However, telomere addition to DSBs does not seem compatible with retention of *VSG221*, which in theory would be lost if this repair mechanism were used in HRES.

## 4.6 Analysis of cell cycle, VSG221 expression and H2Ax nuclear signal after DSB induction in *pif6* mutants

### 4.6.1 Analysis of cell cycle arrest following DSB induction in *HRES* and *HR1 pif6* mutants

DSB induction in HR1 and HRES cells has been demonstrated to cause cell cycle arrest at G2/M at 12 and 24 hours post tetracycline addition (Glover *et al.*, 2013a; Glover *et al.*, 2008). Given the potentially unusual profile of increased survival after I-SceI induction, the cell cycle profile of HRES wild type and *HRES pif6* mutants was assayed to investigate whether loss of *pif6* affected cell cycle arrest following DSB formation and perhaps, and suggest an explanation for how this helicase might be involved in the response to a DSB.



HRES wild type and *pif6* mutant cells were induced with  $2\ \mu\text{g.mL}^{-1}$  tetracycline and cells were collected at 12 and 24 hours. Cells were stained with DAPI to enable visualisation of nuclear and kinetoplast DNA and cells were categorised as either 1N1K (G1 phase), 1N2K (G2 phase), 2N2K (M phase), or 'other' for cells that were abnormal and corresponded to none of these categories. See Section 3.4.2 for a full explanation of the categorisation of visible cell cycle stages.

The expected distribution of cell types for unperturbed BSF *T. brucei* is ~80% 1N1K, ~10% 1N2K and ~5% 2N2K (Benmerzouga *et al.*, 2013). The cell cycle profile for uninduced HRES wild type and *HRES pif6*<sup>+/-</sup> cells was consistent with the expected distribution (Fig. 4-18A&B). However, an increase in 1N2K (~G2 phase) cells was observed in *HRES pif6*<sup>-/-</sup> uninduced cells at both 12 and 24 hours, increasing from ~10% in both HRES wild type and *HRES pif6*<sup>+/-</sup> to ~16%. The proportion of 2N2K cells was also higher in uninduced *HRES pif6*<sup>-/-</sup> cells at 24 hours (~9%) compared to HRES wild type and *HRES pif6*<sup>+/-</sup> cells (both ~5%). However, this was not apparent at the 12 hour time point and it is unclear why this would change without induction of I-SceI. The higher proportion of 1N2K *HRES pif6*<sup>-/-</sup> cells suggests a cell cycle abnormality in *HRES pif6*<sup>-/-</sup> cells involving a delay in progression to M phase.

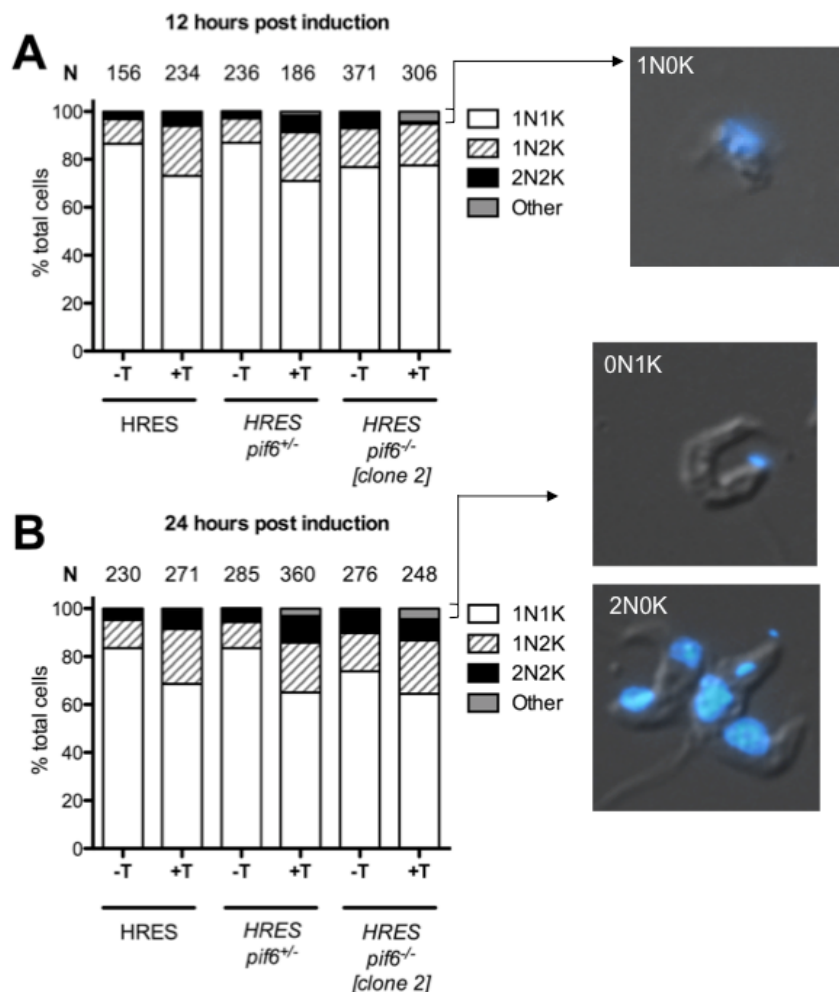
Twelve hours after addition of tetracycline, an increase in 1N2K cells was observed in induced HRES wild type and *HRES pif6*<sup>+/-</sup> cultures, with the proportion of such cells increasing from ~10% (in the uninduced) to ~20%, which was accompanied by a decrease in 1N1K cells from ~80% (uninduced) to 70% (Fig. 4-18A). This is similar to the increase in G2 phase cells (to ~25%) observed by Glover *et al.* (2013a) after I-SceI induction in HRES cells. In contrast, the 1N2K population of induced *HRES pif6*<sup>-/-</sup> cells did not increase above the 16% of uninduced cells at this time point.

24 hours after I-SceI induction (Fig. 4-18B), the proportion of 1N2K cells remained elevated at ~20% in HRES wild type and *HRES pif6*<sup>+/-</sup> cells, accompanied by an increase in the proportion of 2N2K cells (from ~5% to ~10%), indicating a block in cytokinesis. The proportion of 2N2K cells was already elevated in uninduced *HRES pif6*<sup>-/-</sup> cells at 24 hours (~10%) and did not increase above this level. However, the proportion of 1N2K cells was elevated above the uninduced level (~16%), though only modestly (to ~22%).

One other feature was potentially seen following I-SceI induction: a small but noticeable increase in the proportion of ‘other’ cells was observed in *HRES pif6<sup>-/-</sup>* mutants. Such aberrant cells were not seen at all in any uninduced HRES populations and nor in the induced wild type, but were seen to a very limited extent in the *pif6<sup>+/-</sup>* cells, where their number appeared to increase from 12 to 24 hours, though never reaching the levels in the *pif6<sup>-/-</sup>* mutants. The ‘others’ were of several different categories, including zoid cells (0N1K), cells without kinetoplasts (e.g. 2N0K and 1N0K) and cells containing a fragmented nucleus that appeared to have begun to break down. Examples are shown in Figure 4-18. Due to the low number of ‘other’ cells, their types were not sub-categorised during counting.

The HRES wild type data above are consistent with the findings of Glover *et al.* (2013a), namely, that DSB induction in HRES wild type causes arrest at the G2/M cell cycle checkpoint and a cytokinesis block at 12 hours post-induction. I-SceI induction in *HRES pif6<sup>+/-</sup>* mutants results in the same cell cycle arrest profile as wild type HRES cells, though there may be a small indication that this is impaired, leading to aberrant cells. In contrast, *HRES pif6<sup>-/-</sup>* cells appear to suffer from an intrinsic cell cycle block, reflected in an elevated proportion of 1N2K cells in uninduced cultures, though induction of I-SceI can increase this phenotype (at least to some extent after 24 hrs). Again, the accumulation of aberrant cells after I-SceI induction may indicate some impediment to the cell’s response to DSB induction. The intrinsic cell cycle phenotype (increased 1N2K cells) was not observed in *pif6<sup>-/-</sup>* mutants in a non-I-SceI background (see Section 3.4.2), and it is therefore possible that this is a specific response to DSBs. For instance, expression of the I-SceI meganuclease may be leaky in the HRES cell line. Alternatively, such leakiness may be specific to the clone examined here. It has also been suggested that DSBs occur more frequently at the 70 bp repeats of the active VSG BES (Boothroyd *et al.*, 2009), though this has been contested by later studies (Glover *et al.*, 2013a; Jehi *et al.*, 2014b). Additionally, introduction of the I-SceI/*PUR* construct itself, rather than I-SceI ‘leakiness’, may result in a higher rate of break formation by interfering with transcription or replication for example. Finally, If PIF6 is important in repair of these breaks, this could also cause an increase in 1N2K cells in the absence of I-SceI

induction, as observed here.

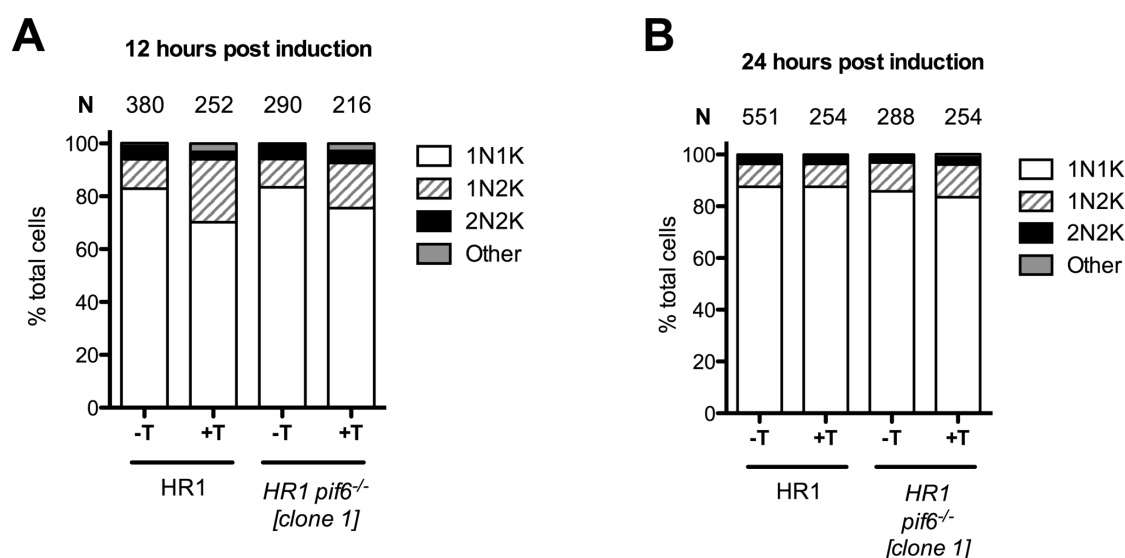


**Figure 4-18 DAPI cell cycle analysis of *HRES pif6* mutants**

The DNA content of *HRES* wild type, *HRES pif6*<sup>+/-</sup> and *HRES pif6*<sup>-/-</sup> cells at 12 hours and 24 hours post I-SceI induction (+T) was visualised by DAPI staining of fixed cells. The number of cells with one nucleus and one kinetoplast (1N1K), one nucleus and two kinetoplasts (1N2K), two nuclei and two kinetoplasts (2N2K) and cells that did not fit into any of these categories (other) were counted using captured images. The proportions of each of these cell types are represented as a percentage of the total cells counted (N). Uninduced cells (-T) were also analysed as a control. N, number of cells counted. Examples of 'other' cells from the cell lines and times indicated are shown.

To test the above data further, the cell cycle profile of *HR1* wild type and *HR1 pif6*<sup>-/-</sup> mutants was analysed (Fig. 4-19A&B). In the absence of I-SceI induction, the cell cycle profile was normal for *HR1* wild type and *HR1 pif6*<sup>-/-</sup>. However, induction of I-SceI resulted in the accumulation of 1N2K cells in *HR1* wild type and *HR1 pif6*<sup>-/-</sup> at 12 hours (Fig. 4-19A), indicating G2/M cell cycle arrest. Approximately 24% of *HR1* and 17% of *HR1 pif6*<sup>-/-</sup> cells were 1N2K, similar to the accumulation of 1N2K cells found after I-SceI induction in *HRES*. However, the G2/M cell cycle arrest observed in *HR1* wild type and *HR1 pif6*<sup>-/-</sup> resolved more

rapidly than in HRES wild type and *HRES pif6* mutants, with the proportion of 1N2K cells returning to normal in induced cultures (~10%) at 24 hours (Fig. 4-19B), perhaps suggesting that a DSB formed in the chromosomal internal location of HR1 is repaired more efficiently than the DSB formed at the 70 bp repeats in HRES. These data are consistent with the G2/M cell cycle arrest and resolution observed by Glover *et al.* (2008) in HR1 following I-SceI induction, and that the altered cell cycle profiles in the *pif6* mutants appear to be specific to HRES cells.



**Figure 4-19** Cell cycle analysis of *HR1 pif6* mutants following I-SceI induction. The DNA content of HRES wild type and *HRES pif6*<sup>-/-</sup> cells at 12 hours and 24 hours post I-SceI induction (+T) was visualised by DAPI staining of fixed cells. The number of cells with one nucleus and one kinetoplast (1N1K), one nucleus and two kinetoplasts (1N2K), two nuclei and two kinetoplasts (2N2K) and cells that did not fit into any of these categories (other) were counted using captured images. The proportions of each of these cell types are represented as a percentage of the total cells counted (N). Uninduced cells (-T) were also analysed as a control. N, number of cells counted.

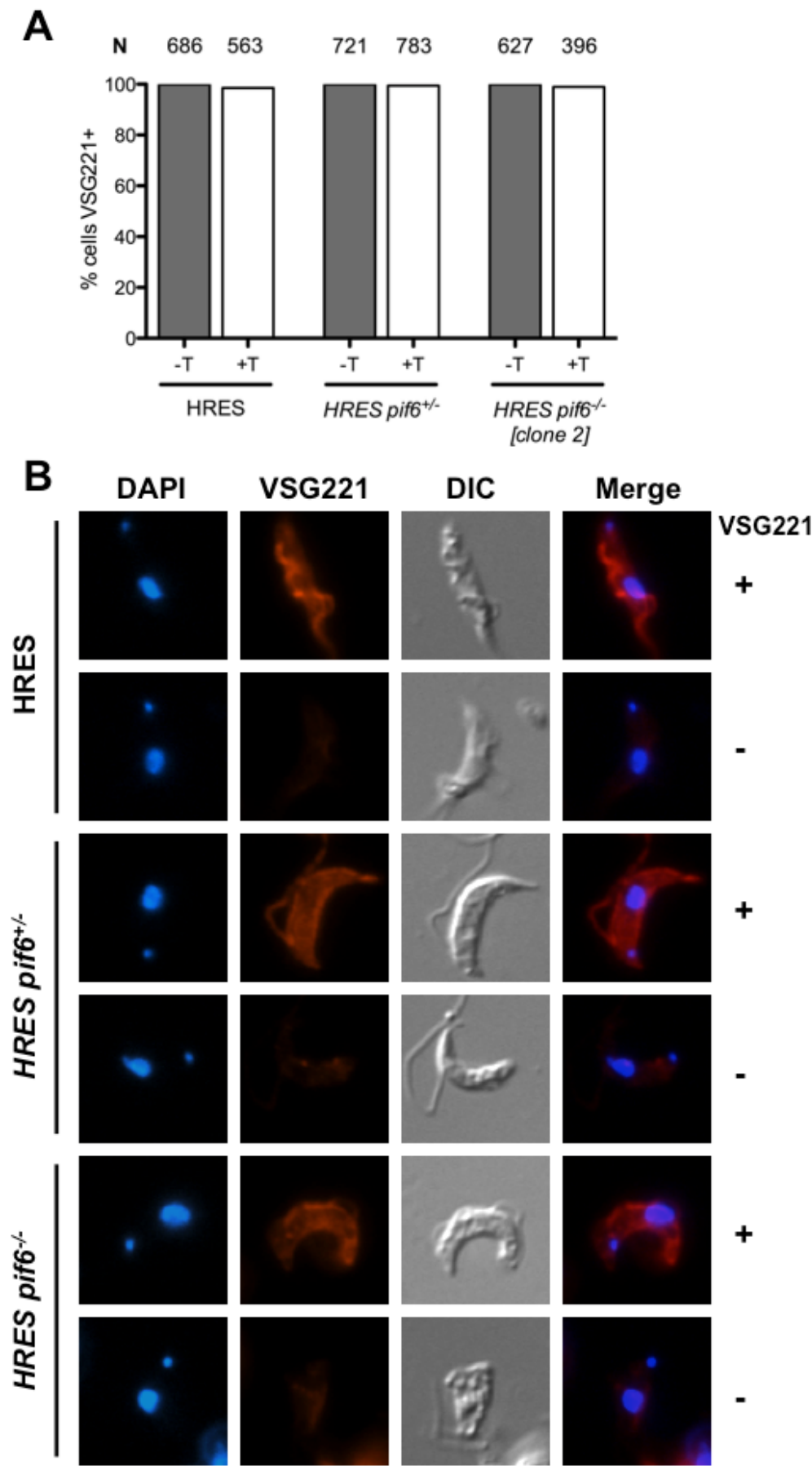
#### 4.6.2 Analysis of VSG221 expression after DSB induction in *HRES pif6* mutants

To investigate whether loss of *PIF6* affects loss of VSG221 expression following DSB induction in an HRES background, as might be expected if a DSB drives VSG switching, immunolocalisation of VSG221 following I-SceI induction was carried out. Cells were induced with 2  $\mu\text{g.mL}^{-1}$  tetracycline and samples were collected for VSG221 immunolocalisation at 24 hours. Rabbit anti-VSG221 antiserum (Glover *et al.*, 2013a) was used at a 1:1000 dilution and a secondary anti-rabbit Alexa Fluor 594 conjugated antiserum (Life Technologies) was used at a 1:2000 dilution to detect VSG221. From previous work, the accumulation of 1N2K (G2)

cells has been shown to peak at 12 hours after I-SceI induction in HRES wild type cells, then return to uninduced levels by 24 hours (Glover *et al.*, 2013a; Glover & Horn, 2014). It was therefore expected that at 24 hours post-induction the proportion of VSG221-positive cells would decrease, due to an induced switch in expression to a different VSG. However, Figure 4-20 shows that any reduction in the proportion of cells expressing VSG221 was negligible at this stage after induction, with only 0.5-1% of cells not displaying VSG221 signal. It is possible that this is because in these experiments cell cycle arrest following DSB induction was still present 24 hours after induction (discussed above, Section 4.6.1), in contrast to the reports by Glover and colleagues. Why this arrest should be more prolonged here is not clear, but given a prolonged cell cycle arrest, it is perhaps not surprising that VSG221 on the cell surface has not been replaced by another VSG. To explore whether loss of *PIF6* affects loss of VSG221 expression following DSB induction, cell samples would need be collected at time points beyond 24 hours in order to examine cells that have re-entered the cell cycle. Alternatively, clonal dilution of surviving clones ~one week after induction would provide data on whether VSG221 expression is retained in *pif6*<sup>-/-</sup> mutants, since virtually all clonal wild type HRES survivors do not express VSG221 (Glover *et al.*, 2013a; Glover & Horn, 2014). However, the timing of a switch to expression of another VSG could not be examined using a clonal survival method.

**Figure 4-20 Analysis of VSG221 expression in HRES *pif6* mutants following I-SceI induction**

VSG221 expression in HRES wild type, HRES *pif6*<sup>+/+</sup> and HRES *pif6*<sup>-/-</sup> cells at 24 hours post I-SceI induction (+T) was visualised by staining fixed cells with rabbit anti-VSG221 antiserum (1:1000 dilution, (Glover *et al.*, 2013a)) and secondary staining with goat anti-rabbit Alexa Fluor 594-conjugated antiserum (1:2000, Life Technologies). The proportion of cells expressing VSG221 on their surface at 24 hours post I-SceI induction (+T) or without I-SceI induction (-T) is represented as the percentage of total cells counted. N, total number of cells counted.

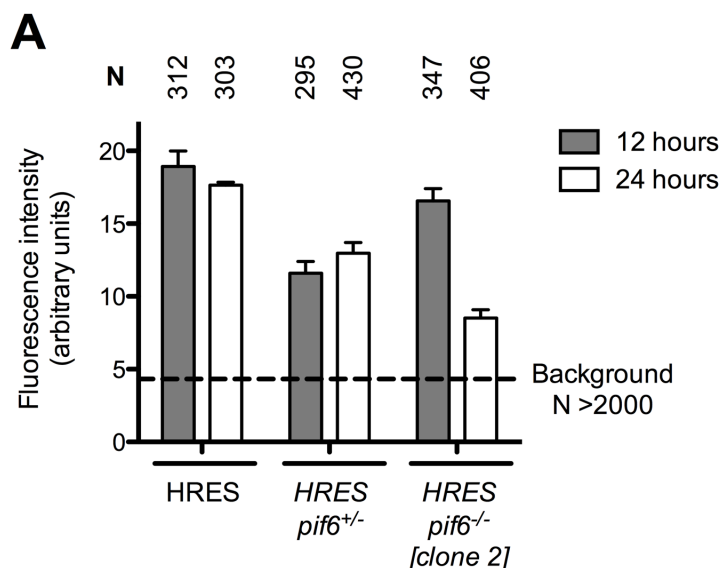


### 4.6.3 Analysis of H2AX nuclear signal after DSB induction in *HRES* and *HR1 pif6* mutants

One of the earliest markers of DNA damage in eukaryotes is  $\gamma$ H2A(X). In mammals, the variant histone H2AX is Ser<sup>139</sup> phosphorylated (Rogakou *et al.*, 1998) and in *S. cerevisiae* (Downs *et al.*, 2000; Redon *et al.*, 2003) and *S. pombe* (Nakamura *et al.*, 2004),  $\gamma$ H2A is a Ser<sup>128/129</sup> phosphorylated version of the core histone H2A.  $\gamma$ H2A(X), from here on referred to as  $\gamma$ H2A, signals DNA damage and triggers the recruitment of DNA repair proteins (Canman, 2003; Polo & Jackson, 2011) including MDC1 (mediator of damage checkpoint 1) (Stucki *et al.*, 2005), which activates the G2/M and intra S-phase cell cycle checkpoints (Stucki & Jackson, 2006). The *T. brucei* equivalent of  $\gamma$ H2A has been identified and characterised by Glover *et al.* (2012), who showed that histone H2A is not phosphorylated on a C-terminal serine residue, but instead that Thr<sup>130</sup> is phosphorylated to form  $\gamma$ H2A, whose abundance increases after DNA damage, including at sites of I-SceI-induced DSBs, and becomes visible as subnuclear foci.

To investigate whether loss of PIF6 affects accumulation of  $\gamma$ H2A in response to DSB formation, the intensity of  $\gamma$ H2A immunofluorescence was measured in wild type and *pif6* mutant HR1 and HRES cells following I-SceI induction. Cultures were induced with 2  $\mu\text{g.mL}^{-1}$  tetracycline and cells collected at 12 and 24 hours post-induction. Cells were fixed and stained for  $\gamma$ H2A using rabbit anti- $\gamma$ H2A antiserum at a 1:1000 dilution (gift, Tiago D. Serafim, unpublished) and goat anti-rabbit Alexa Fluor 594-conjugated antiserum (Life Technologies) at a 1:2000 dilution and slides were also stained with DAPI for visualisation of the nucleus and kinetoplast. Although  $\gamma$ H2A in trypanosomes has previously been analysed by looking at nuclear  $\gamma$ H2A foci (Glover & Horn, 2012), the antiserum used here did not provide clear results using this method. The staining obtained was general nuclear staining in the vast majority of cases, not  $\gamma$ H2A foci (see examples in Figure 4-21B). It was therefore decided to measure the nuclear  $\gamma$ H2A intensity instead. The intensity of  $\gamma$ H2A signal in the nucleus was calculated by measuring the average intensity of all nuclei in captured images using ImageJ ([imagej.nih.gov/ij](http://imagej.nih.gov/ij)) (Rasband, 1997-2014) and is presented in arbitrary units.

The average  $\gamma$ H2A nuclear intensity in all HRES wild type and *HRES pif6* mutant cultures at 12 and 24 hours following I-SceI induction was at least twice the background  $\gamma$ H2A nuclear intensity, as determined by the average nuclear  $\gamma$ H2A intensity of I-SceI uninduced cells (Fig. 4-21A). All fluorescence intensity values cited below are the value above background fluorescence, taking the background value as zero. At 12 hours following I-SceI induction the nuclear  $\gamma$ H2A intensity was highest in HRES wild type (14.6 units above background) and remained almost unchanged (13.4 units above background) at 24 hours. *HRES pif6*<sup>+/-</sup>  $\gamma$ H2A intensity at 12 hours post-induction was ~50% lower (7.3 units above background) than HRES wild type and similarly did not decrease at 24 hours. Therefore, although the  $\gamma$ H2A intensity was lower in *HRES pif6*<sup>+/-</sup> than HRES wild type, neither underwent a substantial change in fluorescence between 12 and 24 hours. The  $\gamma$ H2A intensity profile of *HRES pif6*<sup>-/-</sup> cells however was very different. The  $\gamma$ H2A intensity of *HRES pif6*<sup>-/-</sup> at 12 hours post-induction (12.3 units above background) was comparable to that of wild type HRES. However, at 24 hours post-induction there was a ~60% reduction in *HRES pif6*<sup>-/-</sup>  $\gamma$ H2A intensity (5 units), a signal loss that was not seen in the two other cell lines.



**Figure 4-21 Analysis of  $\gamma$ H2A expression in *HRES pif6* mutants**  
 $\gamma$ H2A expression in HRES wild type, *HRES pif6*<sup>+/-</sup> and *HRES pif6*<sup>-/-</sup> cells at 12 and 24 hours post I-SceI induction was visualised by staining fixed cells with rabbit anti- $\gamma$ H2A antiserum (1:1000 dilution, gift, Tiago D Serafim (unpublished)) and secondary staining with goat anti-rabbit Alexa Fluor 594-conjugated antiserum (1:2000, Life Technologies). (A) The average nuclear intensity of  $\gamma$ H2A expression was measured from captured images using ImageJ ([imagej.nih.gov/ij](http://imagej.nih.gov/ij)) (Rasband, 1997-2014) and is expressed as the average nuclear intensity (arbitrary units) of all nuclei counted (N). The average nuclear intensity of I-SceI uninduced cells was used as a background level (dashed line). Bars represent standard error of the mean (SEM). N, number of cells counted. (B) Representative images of  $\gamma$ H2A expression in



cells at 12 and 24 hours post induction. -T, no tetracycline (control); +T, tetracycline added. Scale bar, 13  $\mu$ M.

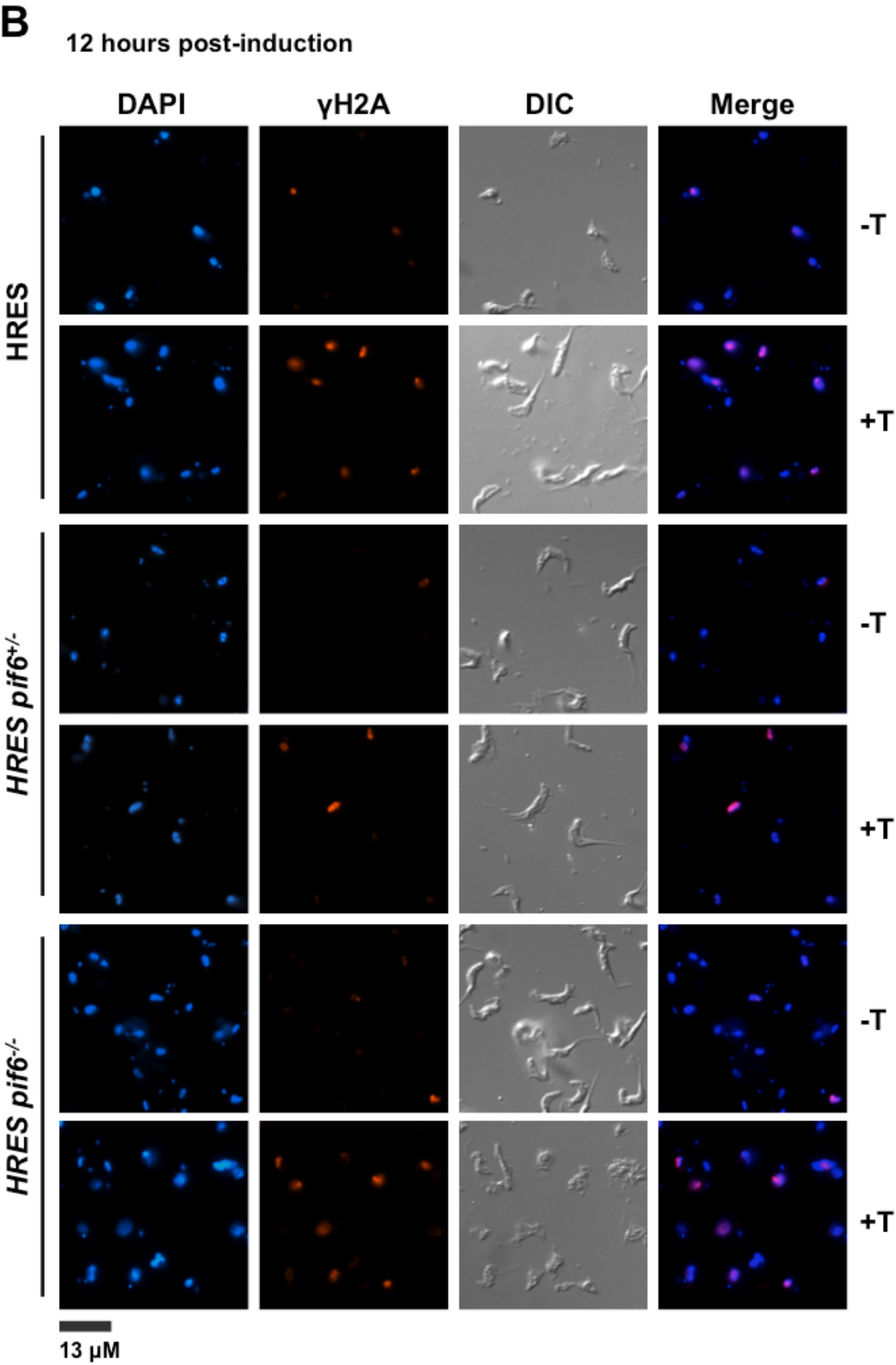


Figure 4-21 continued

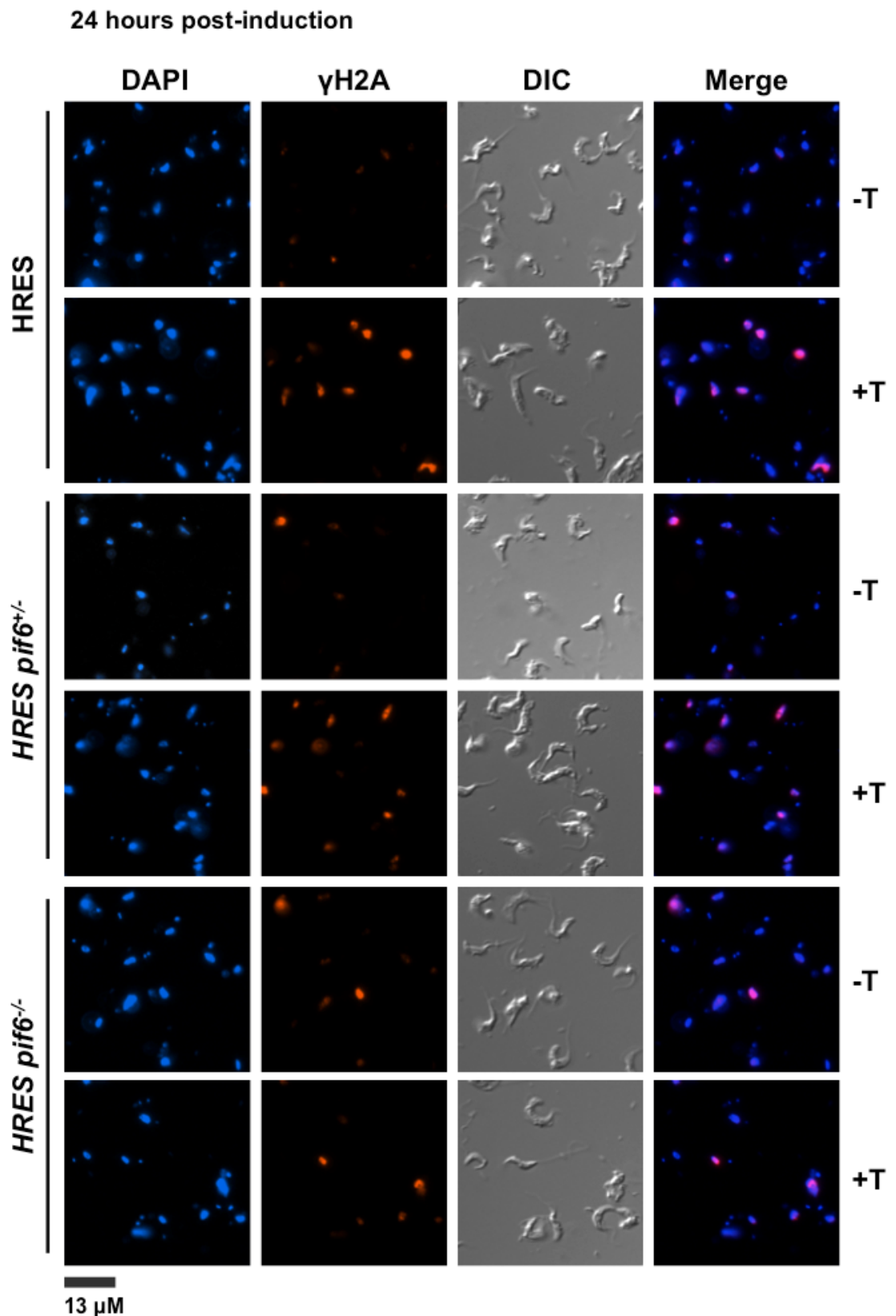
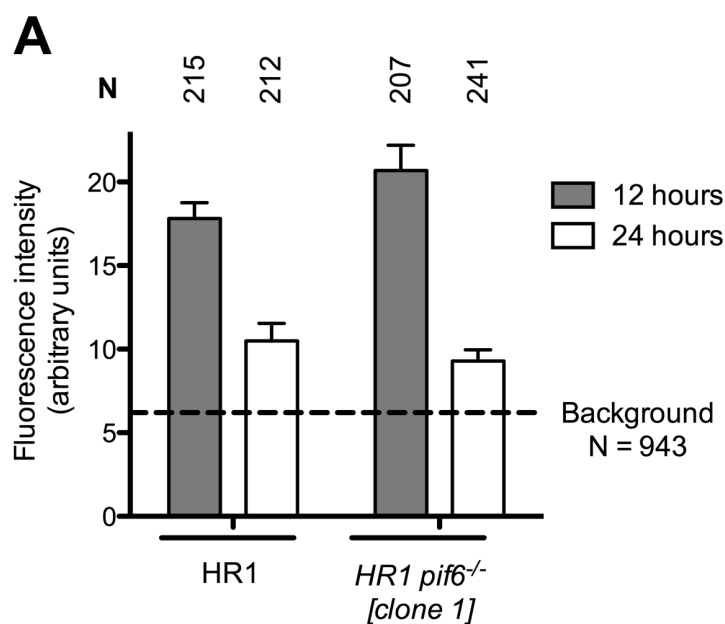


Figure 4-21 continued

Glover *et al.* (2012) showed that  $\gamma$ H2A foci appeared most frequently in S-phase and G2-phase cells. Glover *et al.* (2013a & 2008) also found that I-SceI induction

in HRES cells results in a G2/M phase arrest, which peaks at 12 hours and is resolved by 24 hours (see Section 1.6.1). The data here are largely consistent with this, with the exception that in our hands the G2/M arrest is more slowly resolved for reasons that are unclear. If so, the  $\gamma$ H2A intensity of HRES wild type and *HRES pif6<sup>+/-</sup>* may follow I-SceI induction. At 24 hours the accumulation of 1N2K cells is still present at the same level, and thus it is likely that  $\gamma$ H2A accumulation is still present at this stage; note that the timing of  $\gamma$ H2A focal accumulation and resolution relative to G2/M arrest has not been reported for HRES. However, there was also an elevated level of G2/M arrested *HRES pif6<sup>-/-</sup>* cells at 24 hours post induction. It might therefore be expected that  $\gamma$ H2A intensity would remain similarly elevated at 24 hours in *HRES pif6<sup>-/-</sup>*. Recovery from cell cycle arrest and re-entry into the cell cycle in *S. cerevisiae* is contingent upon dephosphorylation of H2AX (Bazzi *et al.*, 2010; Keogh *et al.*, 2006; Nakada *et al.*, 2008). The lowered levels of  $\gamma$ H2A staining at 24 hours might therefore suggest that *HRES pif6<sup>-/-</sup>* cells undergo H2AX dephosphorylation earlier than HRES wild type and *HRES pif6<sup>+/-</sup>* cells, but have not yet resumed the cell cycle at the time point examined.

The nuclear  $\gamma$ H2A signal intensity of HR1 wild type and *HR1 pif6<sup>-/-</sup>* cells was calculated in the same way as for HRES cells and cells were prepared identically. All values quoted are average fluorescence above the background level. As in the HRES background (see Section 4.6.3), the  $\gamma$ H2A signal intensity of HR1 wild type cells 12 hours post I-SceI induction (11.6 units) was similar to that of *HR1 pif6<sup>-/-</sup>* (14.5 units) (Fig. 4-22). However, whereas the  $\gamma$ H2A signal remained high in HRES wild type cells at 24 hours post induction, the  $\gamma$ H2A signal decreased substantially in both HR1 wild type and *HR1 pif6<sup>-/-</sup>* at 24 hours post I-SceI induction (~60% and ~80% respectively). These data are consistent with the accumulation of 1N2K cells in HR1 wild type and *HR1 pif6<sup>-/-</sup>* 12 hours post I-SceI induction and the return to a normal cell cycle profile of both at 24 hours post-induction (see Section 4.6.1). This indicates that DSBs at the chromosomal internal location in HR1 cells are repaired more efficiently than DSBs at the 70 bp repeats of the active VSG BES of HRES cells. Furthermore, H2A phosphorylation following I-SceI induction in HR1, an important marker of DNA damage, is resolved as efficiently in *HR1 pif6<sup>-/-</sup>* cells as HR1 wild type cells.



**Figure 4-22 Analysis of  $\gamma$ H2A expression in *HR1 pif6* mutants**  
 $\gamma$ H2A expression of HR1 wild type and *HR1 pif6*<sup>-/-</sup> cells at 12 and 24 hours post I-SceI induction was visualised by staining fixed cells with rabbit anti-  $\gamma$ H2A antiserum (1:1000 dilution, gift, Tiago D Serafim (unpublished)) and secondary staining with goat anti-rabbit Alexa Fluor 594-conjugated antiserum (1:2000, Life Technologies). (A) The average nuclear intensity of  $\gamma$ H2A expression was measured from captured images using ImageJ ([imagej.nih.gov/ij](http://imagej.nih.gov/ij)) (Rasband, 1997-2014) and is expressed as the average nuclear intensity (arbitrary units) of all nuclei counted (N). The average nuclear intensity of I-SceI uninduced cells was used as a background level (dashed line). Bars represent standard error of the mean (SEM). (B) Representative images of  $\gamma$ H2A expression in cells at 12 and 24 hours post induction. -T, no tetracycline (control); +T, tetracycline added. Scale bar, 13  $\mu$ M.

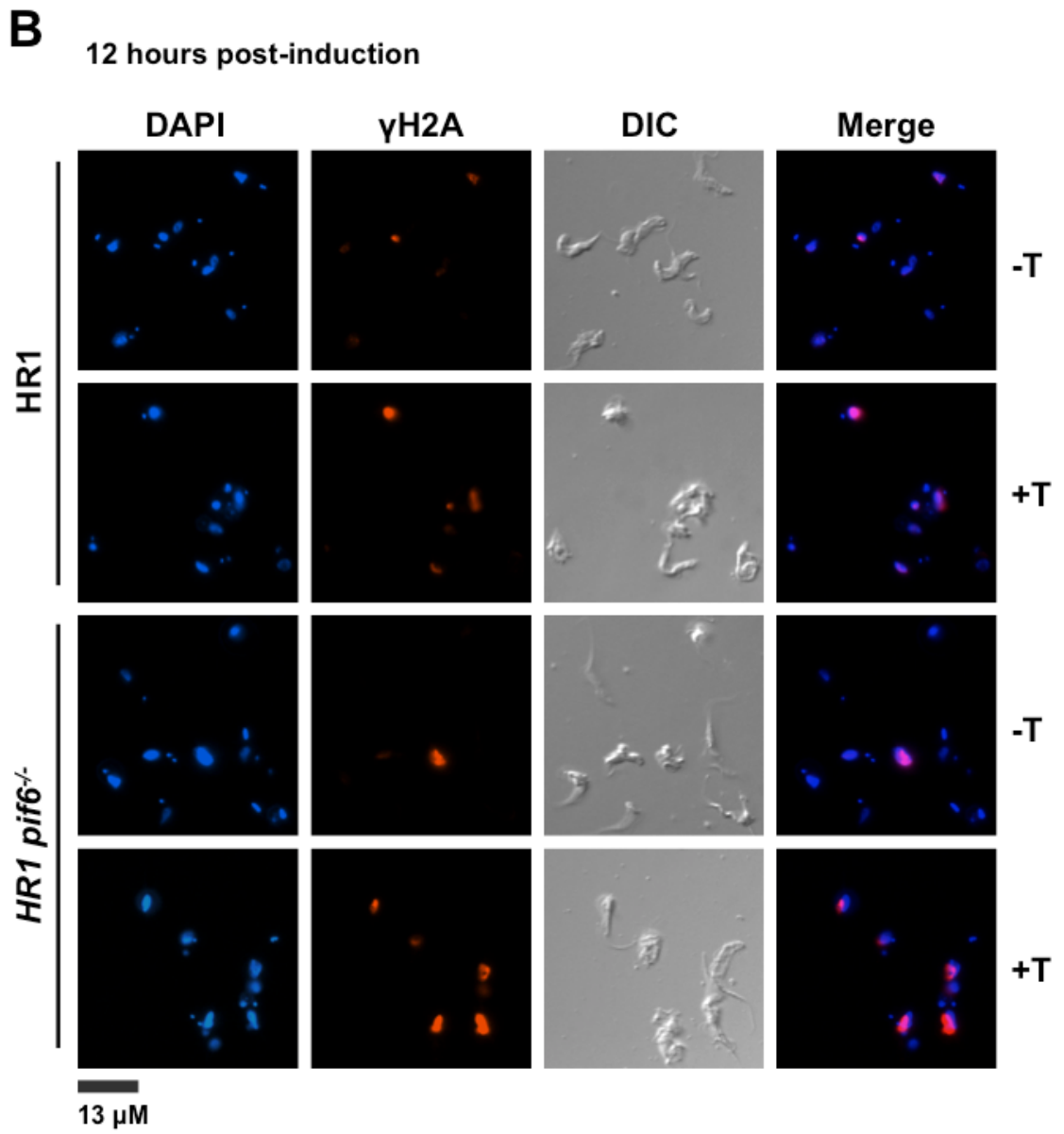


Figure 4-22 continued

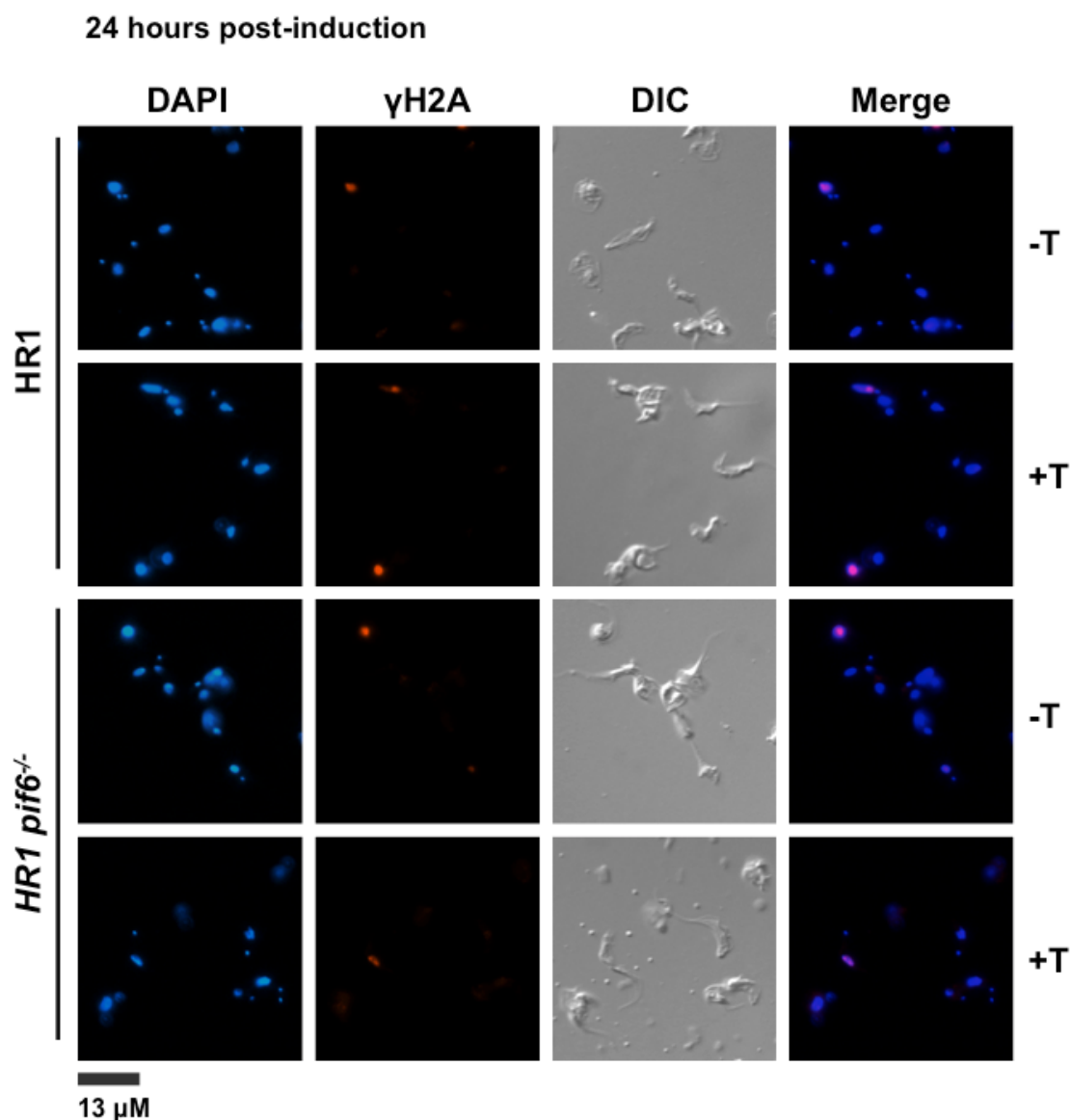


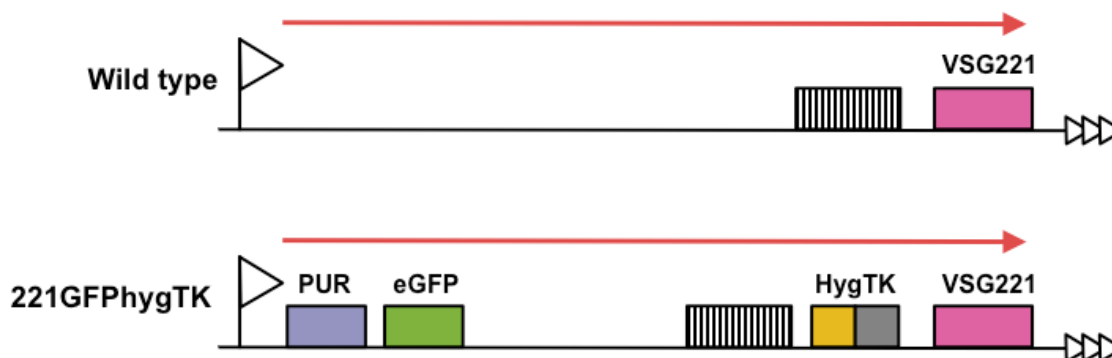
Figure 4-22 continued

## 4.7 Analysis of VSG switching in knockout cell lines

### 4.7.1 VSG switching assay strategy

In order to analyse the effect of loss of *RECQ2*, *MUS81* and *PIF6* on VSG switching *in vitro* we adapted the strategy of Povelones *et al.* (2012), who used RNAi against histone H1 to examine the role of this factor in VSG switching. In this strategy, a herpes simplex virus thymidine kinase (*HSV-TK*, here referred to as *TK*) gene is fused to a hygromycin resistance gene (*HYG-TK*) and inserted between the 70 bp repeats and *VSG221* (427-2) in the active VSG BES (named BES1, as described by directed cloning in Hertz-Fowler *et al.* (2008) of *T. brucei* Lister 427 strain MITat1.2. Additionally, enhanced *GFP* (*eGFP*) and a puromycin

resistance gene are integrated downstream of the active *VSG* BES promoter (Fig. 4-24). This cell line is here referred to as *GFP221hygTK*.



**Figure 4-23** Cell line *GFP221hygTK* used for VSG switching experiments  
As discussed in the text, the *GFP221hygTK* cell line contains a puromycin resistance gene (*PUR*, purple) and *eGFP* gene (*GFP*, green) integrated immediately downstream of the *VSG221* BES (active *VSG* BES) promoter (arrow) allowing monitoring of the activity status of the expression site, and a hygromycin-TK (*Hyg-TK*, orange-grey) fusion gene integrated between the 70 bp repeats (hatched) and *VSG221* (pink), to permit negative selection of *VSG221* in the active *VSG* BES.

In this experimental approach to analyse VSG switching *GFP221hygTK* cells are maintained under puromycin and hygromycin selection in order to maintain transcription of the *VSG221* BES and presence of *VSG221* in the BES, respectively. Thymidine kinase is widely used as a negative selection marker in mammalian systems: to enrich for correctly integrated DNA insertions (Mansour *et al.*, 1988) and to eliminate certain cell types (Heyman *et al.*, 1989). Thymidine kinase phosphorylates a range of nucleosides and nucleoside analogues, including ganciclovir, into toxic compounds that cause cell death.

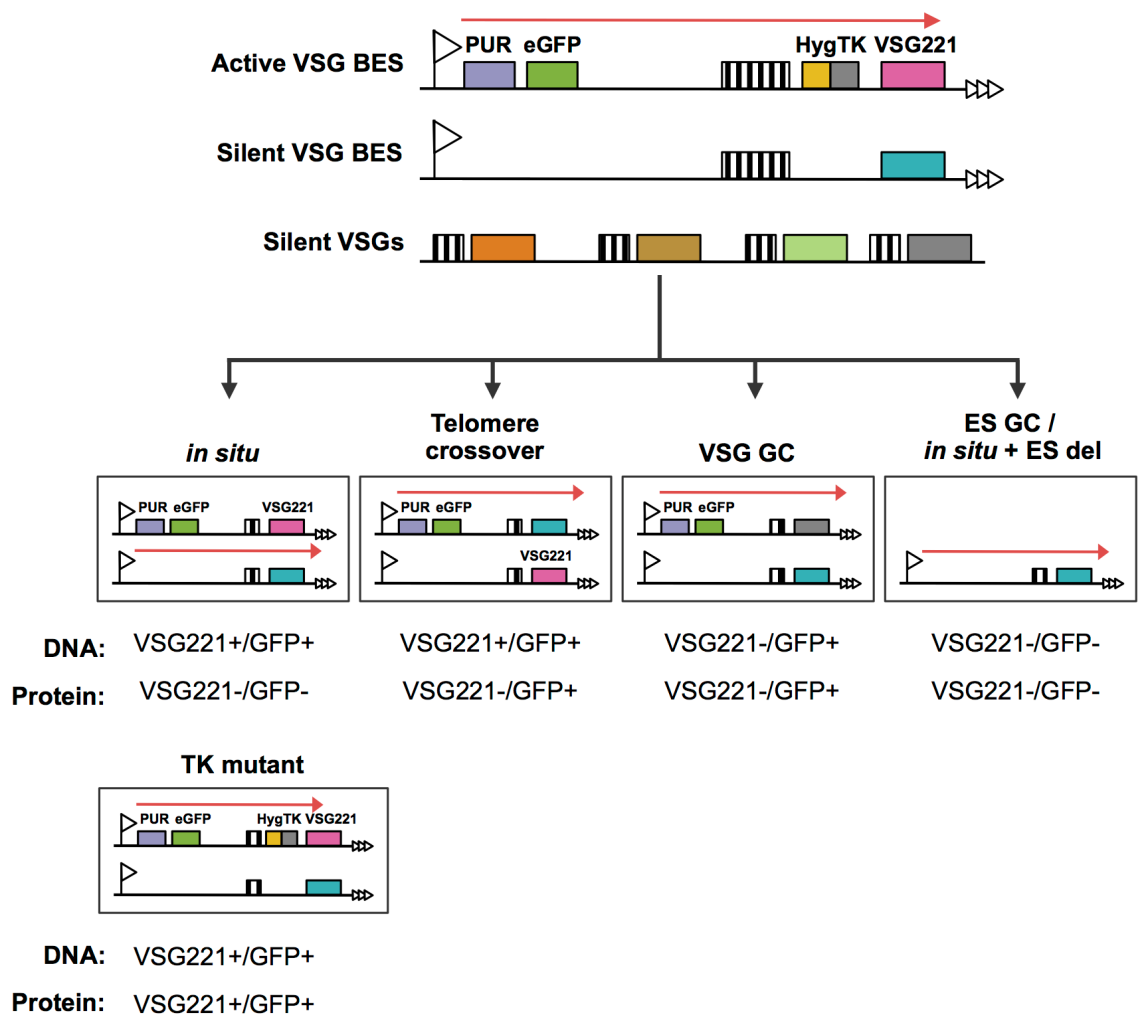
At the beginning of the VSG switching assay cells are passaged into medium not containing hygromycin, either in the presence or absence of puromycin ( $0.2 \mu\text{g} \cdot \text{mL}^{-1}$ ). The absence of hygromycin selection allows cells to lose or inactivate the *HYG-TK* fusion gene upstream of *VSG221* in the active BES. VSG switching can occur by a number of events within the *VSG* BES. Recombination reactions, which are most frequently gene conversions that delete sequences within the *VSG* BES, can involve the region upstream of *VSG221* where the *HYG-TK* fusion gene is located; such recombination can be variable, encompassing merely the *VSG221* proximal region up to the 70 bp repeats, or can involve sequences upstream of the 70 bp repeats, such as the *ESAGs* or upstream of the promoter, in which case the *PUR-GFP* markers would be affected. *VSG* BES

deletion events have also been described. Finally, switching can occur by *in situ* (transcriptional) switching, inactivating the *VSG221* BES and activating another *VSG* BES. Puromycin selection forces cells to maintain transcription of the *VSG221* BES, thus preventing *in situ* switching that would inactivate *PUR-GFP* expression and *VSG* BES deletion *VSG* switching events that remove the *PUR-GFP* markers. After 48 hours of culture in the absence of antibiotic, during which switching can occur, the cells are then diluted in the same media but containing ganciclovir. The assay utilises the toxic activity of TK on ganciclovir to eliminate from the population cells that have not inactivated the TK gene through a *VSG* switching event. Thus, TK negative selection is used a proxy for *VSG* switching. Cells that have inactivated TK through point mutations rather than *VSG* switching will also be selected for in this assay. However, these clones can be identified by their continued expression of *VSG221*, as revealed by western blot analysis.

Analysis of the proportion of clones surviving ganciclovir selection permits the calculation of the *VSG* switching rate. PCR analyses of *VSG221* and *eGFP* presence and western blot analysis of *VSG221* and *eGFP* expression allows the determination of the mechanism of *VSG* switching (Fig. 4-24). *In situ* switchers are positive for *VSG221* and *eGFP* by PCR but no longer express either. Clones that have switched by *VSG* gene conversion downstream of *PUR-GFP* (*VSG* GC) retain *eGFP* in the active expression site and are thus positive for GFP by PCR and western blot analysis, but have lost *VSG221* and are therefore are negative for the *VSG221* gene by PCR and the protein by western blot analysis. Telomere crossover switchers can also retain *eGFP* in the active expression site and thus are also *eGFP* positive by PCR and western blot analysis. However, in this case, *VSG221* is retained in the genome but no longer in the active expression site, and so are *VSG221* positive by PCR but not by western blot analysis. The final types of switch detectable in this system are gene conversions (GC) encompassing the whole expression site or *in situ* switch with expression site deletion (ES GC or *in situ* + ES del, respectively). Both of these switching events result in the loss of the *VSG* BES and are indistinguishable in this assay. They are negative for both *GFP* and *VSG221* by PCR and express neither protein in western analysis.



A similar strategy to the one employed by Povelones *et al.* (2012), and adopted here, has been used previously to investigate VSG switching (Benmerzouga *et al.*, 2013; Jehi *et al.*, 2014a; Jehi *et al.*, 2014b; Kim & Cross, 2010; Kim & Cross, 2011). The strategy used in these previous studies utilised the same principles as this work and Povelones *et al.* (2012), but differed in some technical details: by the absence of *eGFP* at the *VSG221* BES promoter, the use of a blasticidin resistance instead of puromycin resistance cassette at the promoter, as well as the TK-fusion gene at the 70 bp repeat utilising a puromycin resistance cassette instead of hygromycin resistance. It is also likely that differences exist in the precise locations of construct integrations within the *VSG221* BES.



**Figure 4-24 Strategy for determining VSG switching mechanisms**

The active *VSG221hygTK* is shown, within which the *PUR*, *GFP*, *HYG-TK* and *VSG221* genes are represented as in Figure 4-24. In addition, one of the ~15 silent *VSG* BESs containing a distinct *VSG* (turquoise) is shown, as are multiple silent *VSGs* elsewhere in the genome (various colours; for convenience these are shown as an array, but could also be at the telomere of silent mini-chromosomes). Upstream of all *VSGs* 70 bp repeats are denoted by hatched boxes. After ganciclovir treatment, utilisation of different switching strategies results in different outcomes (boxes), distinguishable by analysis of *VSG221* and *GFP* presence by PCR and expression by western blot (profiles detailed under each

mechanism). Switching by *in situ* switching, telomere crossover (XO), VSG gene conversion (VSG GC), and BES gene conversion/*in situ* switch + BES deletion (ES GC/ *in situ* + ES del) are detectable using this strategy. TK mutants can also be identified using these analyses. Note, for VSG GC the silent grey array donor VSG gene is shown as being copied but the reaction could also use a BES VSG gene; in addition, *in situ* + ES del is shown and not ES GC, where all sequence of a silent BES is duplicated and replaces the VSG221 BES. Figure inspired by Povelones *et al.* (2012).

#### 4.7.2 Generation of VSG switching cell line

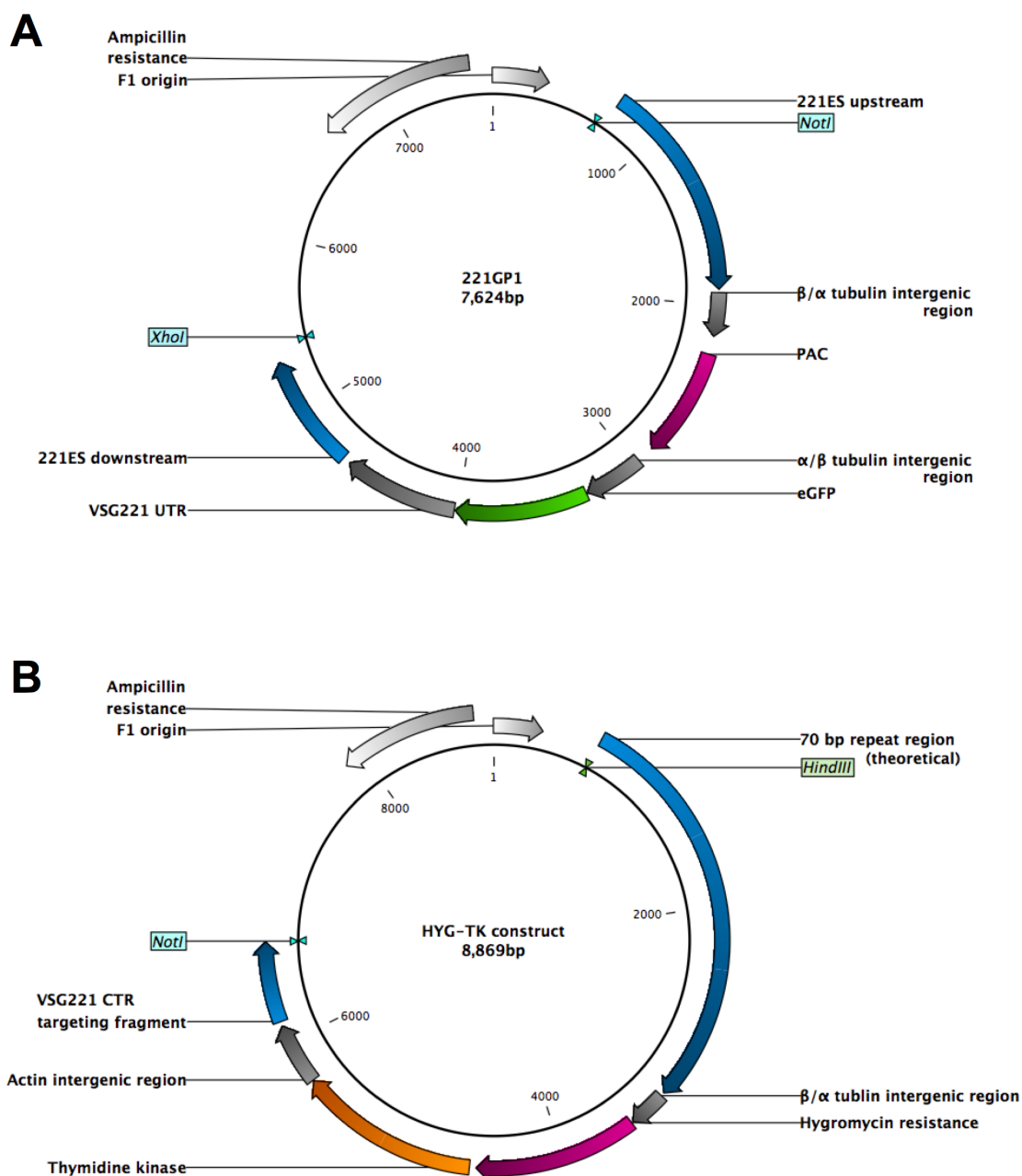
The GFP221hygTK line generated for analysis of VSG switching used the same constructs as in Povelones *et al.* (2012), kindly gifted by Gloria Rudenko. It was necessary to construct the GFP221hygTK strain because the only strain available had been previously genetically modified for RNAi against histone H1. The first of these constructs contained the *eGFP-PUR* (puromycin) construct, 221GP1 (Sheader *et al.*, 2004) (Fig. 4-25A). 221GP1 contains a puromycin resistance gene (*PUR*) flanked by tubulin intergenic sequences, providing splicing and polyadenylation sites. The *eGFP* gene is located downstream of *PUR* and is flanked upstream by the same tubulin intergenic region used downstream of *PUR* and downstream by the VSG221 UTR, providing a polyadenylation signal. The *PUR-eGFP* sequence is flanked by 221 BES targeting fragments and the whole construct is bounded 5' by a NotI restriction site and 3' by an XhoI restriction site. Digestion of the 221GP1 construct with NotI and XhoI prior to transformation releases the targeting construct, which integrates 216 bp downstream of the VSG221 BES transcription start site.

The second construct used to generate the VSG switching line contains the *HYG-TK* gene and integrates downstream of the 70 bp repeats. The construct (gift, G. Rudenko (Povelones *et al.*, 2012)), here referred to as HYG-TK, contains the *HYG-TK* gene flanked upstream by a tubulin intergenic region, providing a splicing signal and downstream by an actin intergenic region, providing for polyadenylation (Fig. 4-25B). This cassette is flanked upstream by a targeting sequence composed of a 2.5 kb region of the 70 bp repeats excised from the construct RM3173 (McCulloch *et al.*, 1997) and downstream by a VSG221 BES targeting fragment, which is composed of the so-called 'co-transposed region' found upstream of all VSGs, in this case from the VSG221 BES between the 70 bp repeats and the VSG221 gene. The construct was digested with NotI and HindIII prior to transformation. Significant difficulty was encountered in attempting to propagate the HYG-TK construct. Restriction digest of plasmid preparations

showed that much of the plasmid had been lost, most likely due to the presence of the 70 bp repeats, which may be unstable in *E. coli*. Propagation of the correct, full length HYG-TK construct was ultimately successful using XL-10 Gold Cells (Stratagene) and ZYM-5 medium. These cells contain a number of mutations in genes involved in DNA rearrangement to improve the stability of unstable clones. Both constructs were checked using restriction enzyme digest prior to use in transformations.

Before beginning generation of the GFP221hygTK cell line, it was confirmed by immunofluorescence that the wild type *T. brucei* Lister 427 MITat1.2 BSF cells expressed VSG221, using an anti-VSG221 antiserum (gift, K. Matthews), data not shown. These cells were then transformed (as described in Section 2.1.3) with 10 µg of the 221GP1 construct and selected using 0.2 µg.mL<sup>-1</sup> puromycin. Four clones were retrieved from this transformation. The successful integration of the 221GP1 construct was confirmed by PCR analysis (“GFP PCR”) using primer #120, which binds in the *eGFP* ORF and primer #121, which binds downstream of the integration site within the *ESAG7* ORF (Tb427.BES40.2) and produces a 2.8 kbp amplified product (Fig.4-26). Primer sequences can be found in Appendix 7.1. Three of these clones were assayed using the above PCR, all of which were positive (data not shown). These three clones were also assayed for *eGFP* and VSG221 expression by immunofluorescence. VSG221 expression was detected using rabbit anti-VSG221 antiserum (Keith Matthews, used at 1:1000 dilution) and an anti-rabbit Alexa Fluor 594 conjugated antiserum (Invitrogen, used at 1:2000 dilution); *eGFP* fluorescence was detected directly using a FITC filter. All three clones were positive for both *eGFP* and VSG221 expression (data not shown).

One of the above eGFP-PUR clones was taken forward for the second round of transformation. The eGFP-PUR cell line was cultured in the presence of 0.2 µg.mL<sup>-1</sup> puromycin in medium lacking thymidine, transformed with 10 µg of the HYG-TK construct and selected using 0.2 µg.mL<sup>-1</sup> puromycin and 10 µg.mL<sup>-1</sup> hygromycin. Two surviving clones were obtained from this transformation. Successful integration of the HYG-TK construct was confirmed by PCR (“HYG-TK PCR”), using primers #131, which binds in the *TK* ORF and #133, which binds in the VSG221 ORF, producing a 2.2 kbp PCR product (data not shown). Figure 4-26 shows the primer binding locations of the *HYG-TK* and *eGFP* PCRs.



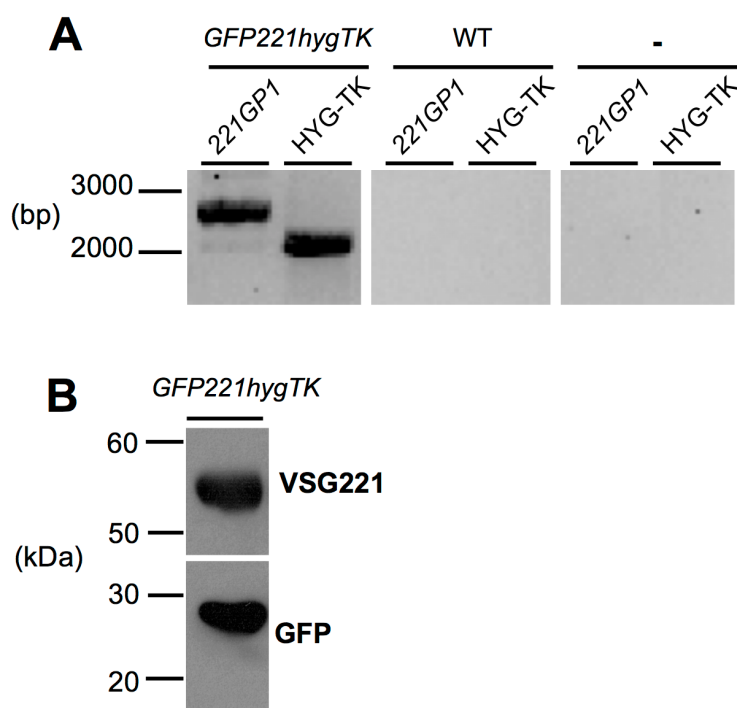
**Figure 4-25** Constructs used in the generation of the GFP221hygTK VSG switching cell line

(A) The 221GP1 construct (Sheader *et al.*, 2004) contains the *eGFP* and *PUR* genes, both flanked by sequences for mRNA splicing and polyadenylation (dark grey). Sequences from the VSG221 BES (221 BES upstream & 221 BES downstream) provide sequence identity for HR with, and integration into, the VSG221 BES immediately downstream of the promoter. Prior to transformation the construct was digested with *XhoI* and *NotI*. (B) The HYG-TK construct (Povelones *et al.*, 2012) contains a hygromycin resistance gene fused to a thymidine kinase gene (*hyg-TK*), flanked by sequences for mRNA splicing and polyadenylation (dark grey). A region of the 70 bp repeats of the VSG221 BES and a region of the VSG221 co-transposed region (VSG221 CTR) provide sequence homology for HR and integration into the VSG221 BES between the 70 bp repeats and VSG221. Prior to transformation the construct was digested with *NotI* and *HindIII*.



**Figure 4-26** PCR primer binding sites of PCR confirm GFP221hygTK cell line  
The active VSG221 BES is shown. Primers #120 and #121 bind in the eGFP and ESAG7 ORFs respectively (“GFP PCR”) to confirm integration of the 221GP1 construct at the correct locus (2.8 kbp). Primers #131 and #133 bind in TK and VSG221 ORFs respectively (“HYG-TK PCR”) to confirm integration of the HYG-TK construct at the correct locus (2.2 kbp product). Primer sequences are listed in Appendix 7.1.

Putative GFP221hygTK clones were also assayed by western blot analysis for eGFP and VSG221 expression and eGFP-PUR construct presence by PCR (“GFP PCR”), as previously. Though both clones were positive for eGFP and VSG221 expression, and 221GP1 and HYG-TK construct integration, only one clone was selected to use to attempt to generate *recq2*, *mus81* and *pif6* mutants. Western blot and PCR confirmation of the selected GFP221hygTK clone is shown in Figure 4-7. Data for the second clone is not shown.



**Figure 4-27** Confirmation of GFP221hygTK cell line by PCR and western blot  
(A) Genomic DNA was extracted from putative GFP221hygTK and wild type cells and integration of the 221GP1 (primers #120 and #121, 2.8 kbp PCR product) and HYG-TK (primers #131 and #133, 2.2 kbp PCR product) constructs was confirmed by PCR amplification. Primer sequences can be found in Appendix 7.1. Distilled water was used as a negative control. Size markers (ladder, bp) are shown. (B) Total protein extract of GFP221hygTK cells were separated by SDS-PAGE, western blotted and membranes were cut in two. The upper half, containing proteins of the predicted molecular weight of VSG221 (51 kDa), was probed with rabbit anti-VSG221 antiserum (1:20,000 dilution, (Glover *et al.*, 2013a)) and the other half, containing proteins of the predicted molecular weight of eGFP (27 kDa) was probed with rabbit anti-GFP antiserum (1:5000 dilution, abcam). Both membranes were probed with goat HRP-conjugated anti-rabbit antiserum (Life

Technologies). Size markers (kDa) are shown. Gaps indicate lanes aligned in this figure after excision from multiple gels/membranes or from disparate parts of the same gel/membrane.

After cells had been transformed with the HYG-TK construct, the resulting GFP221hygTK cells were maintained in culture for as little time as possible, with hygromycin and puromycin antibiotic selection, and always in thymidine-free medium. This was done in order to attempt to minimise the number of cells acquiring inactivating mutations in the thymidine kinase gene, as it has previously been shown that *T. brucei* bloodstream form thymidine kinase expressing cells revert to a ganciclovir-resistant phenotype at a rate of  $10^{-6}$  cells per generation (Valdes *et al.*, 1996). Additionally, the VSG switching strategy employed here has been demonstrated to be prone to the TK mutants (Povelones *et al.*, 2012).

#### 4.7.3 Generation and confirmation of mutants in VSG switching cell line

The plasmids to generate *recq2*, *mus81* and *pif6* mutants in the GFP221hygTK background were the same ones used for the generation of mutants in wild type (Section 3.3.2) and I-SceI (Section 4.3.1) backgrounds. Detailed explanation for the cloning procedure of the knockout plasmids is in Section 3.3.1.

GFP221hygTK cells were cultured in thymidine-free medium containing puromycin ( $0.2 \mu\text{g}.\text{mL}^{-1}$ ) and hygromycin ( $10 \mu\text{g}.\text{mL}^{-1}$ ) and transformed with  $\Delta\text{RECQ2}::\text{BSD}$ ,  $\Delta\text{MUS81}::\text{BSD}$  or  $\Delta\text{PIF6}::\text{BSD}$  in an attempt to generate heterozygous (+/-) cell lines. In all cases, antibiotic resistant clones were selected using  $10 \mu\text{g}.\text{mL}^{-1}$  blasticidin as well as  $5 \mu\text{g}.\text{mL}^{-1}$  hygromycin.

Hygromycin was used to prevent loss of the HYG-TK construct, which was considered a possibility due to the proximity of the integrated construct to the 70 bp repeats and the suggestion that cells containing the HYG-TK construct can lose it during culture (G. Rudenko, personal communication), a suggestion that appeared to be confirmed upon further transformation (see below).

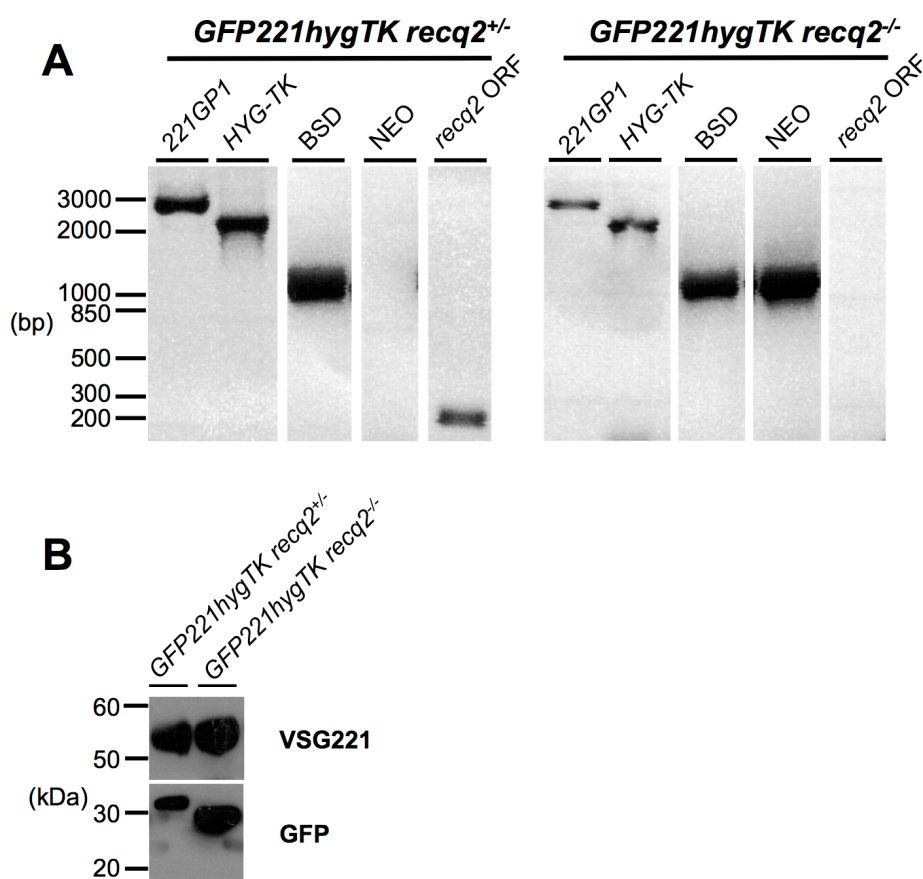
Six putative *recq2*<sup>+/-</sup> (blasticidin resistant) clones were obtained, all of which were taken forward for analysis to confirm correct integration of the construct. A total of 29 blasticidin-resistant positive wells were obtained from the  $\Delta\text{MUS81}::\text{BSD}$  transfection. As 22 of these were on one 24 well plate, and thus were unlikely to be clonal, they were not taken forward for analysis. Six of the

remaining wells (putative *mus81*<sup>+/-</sup> clones, blasticidin resistant) were taken forward for further analysis. Thirteen putative *pif6*<sup>+/-</sup> clones (blasticidin resistant) were obtained from the  $\Delta$ *PIF6::BSD* transformation, of which six were taken forward for analysis.

Putative heterozygote clones were analysed by PCR (data not shown) similar to knockout cells generated in wild type and I-SceI backgrounds (see Sections 3.3.3 and 4.3.2). Briefly, a ~200 bp region of the open reading frame was PCR amplified to test for presence of the ORF of the targeted gene. To check for integration of the blasticidin knockout construct a ~900 bp region was amplified using a forward primer lying upstream of the 5' UTR region present in the knockout construct and a reverse primer specific to the *BSD* gene. To check correct integration of the *NEO* knockout construct in subsequent analysis of putative knockout clones, a ~900 bp region was amplified using a forward primer lying upstream of the 5' UTR region present in the knockout construct and a reverse primer specific to the *NEO* gene. Gene-specific PCR details are listed in the legend of Figures 4-28, 4-29 and 4-30. These resistance gene PCRs were designed to confirm the presence and correct integration location of the knockout cassette relative to the target gene. Sequences of all primers can be found in Appendix 7.1.

In addition to the above PCRs to check correct integration of the knockout construct, the “GFP PCR” and “HYG-TK PCR” (see Section 4.8.2) were performed to confirm the cell lines retained the 221GP1 and HYG-TK constructs, respectively. Expression of *eGFP* and *VSG221* was also assayed by immunofluorescence (as described above in Section 4.8.2). These analyses showed all six putative *GFP221hygTK recq2*<sup>+/-</sup> clones to be correct (data not shown). Four out of the six putative *GFP221hygTK mus81*<sup>+/-</sup> clones were correct, with PCR analyses indicating two clones did not contain the HYG-TK construct and one also lacked the 221GP1 construct (data not shown). Three out of the six putative *GFP221hygTK pif6*<sup>+/-</sup> clones were correct; PCR analyses indicated that two clones lacked the HYG-TK construct and one clone did not contain the correctly integrated  $\Delta$ *PIF6::BSD* cassette (data not shown). One correct clone of each of *GFP221hygTK recq2*<sup>+/-</sup>, *GFP221hygTK mus81*<sup>+/-</sup> and *GFP221hygTK pif6*<sup>+/-</sup> was selected to use in subsequent experiments and for generation of knockout (-/-) clones.

To generate knockout mutants, *GFP221hygTK recq2<sup>+/-</sup>*, *mus81<sup>+/-</sup>* and *pif6<sup>+/-</sup>* cells were transformed with  $\Delta RECQ2::NEO$ ,  $\Delta MUS81::NEO$  or  $\Delta PIF6::NEO$  constructs, respectively. Transformation of *GFP221hygTK recq2<sup>+/-</sup>* and antibiotic selection with  $10 \mu\text{g.mL}^{-1}$  hygromycin,  $5 \mu\text{g.mL}^{-1}$  blasticidin and  $2.5 \mu\text{g.mL}^{-1}$  G418 resulted in one surviving putative *GFP221hygTK recq2<sup>-/-</sup>* clone. Successful amplification of the “NEO”, “BSD”, “HYG-TK” and “GFP” PCR products in the putative *GFP221hygTK recq2<sup>-/-</sup>* clone (Fig. 4-29A) confirmed that the  $\Delta RECQ2::NEO$  and  $\Delta RECQ2::BSD$  cassettes were integrated correctly relative to *RECQ2* and that the clone contained the HYG-TK and 221GP1 constructs. Furthermore, the *RECQ2* ORF could no longer be PCR-amplified, indicating successful knockout of *RECQ2*. Additionally, western blot analysis (Fig. 4-28B) of VSG221 and eGFP expression confirmed that the *GFP221hygTK recq2<sup>-/-</sup>* clone still expressed VSG221 and eGFP.

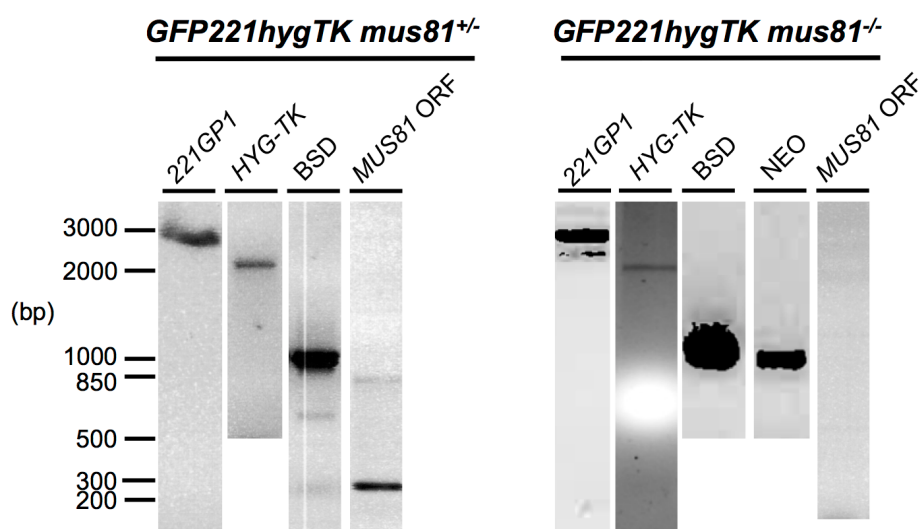


**Figure 4-28** PCR and western blot analysis of *GFP221hygTK recq2* mutants  
**(A)** Genomic DNA was extracted from *GFP221hygTK recq2<sup>+/-</sup>* and *GFP221hygTK recq2<sup>-/-</sup>* cells. A 232 bp region of the ORF (primers #77 & #78) was PCR amplified (“*RECQ2* ORF”) to test for presence of the *RECQ2* ORF. PCR amplification using primers #51 & #154 (1000 bp PCR product) that bind in the *BSD* resistance cassette and upstream of *RECQ2* (“*BSD*”) was used to test for correct integration of the *BSD* knockout construct. PCR amplification using primers #51 and #155 (1008 bp PCR product) that bind in the *NEO* resistance cassette and upstream of *RECQ2* (“*NEO*”) was used to test for correct integration of the *NEO* knockout construct. Presence of the 221GP1 construct was tested by PCR amplification using primers #120 and #120 (2.8 kbp PCR product, “221GP1”) and presence of the HYG-TK



construct was tested by PCR amplification using primers #131 and #133 (2.2 kbp PCR product, “HYG-TK”) constructs. Primer sequences can be found in Appendix 7.1. Size markers are shown (ladder, bp). (B) Total protein extracts were separated by SDS-PAGE and western blotted. Membranes were probed with rabbit anti-VSG221 antiserum (1:20,000 dilution, (Glover *et al.*, 2013a)) and rabbit anti-GFP antiserum (1:5000 dilution, abcam). Both membranes were probed with goat HRP-conjugated anti-rabbit antiserum (Life Technologies). Size markers (kDa) are shown. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels/membranes or from disparate parts of the same gel/membrane.

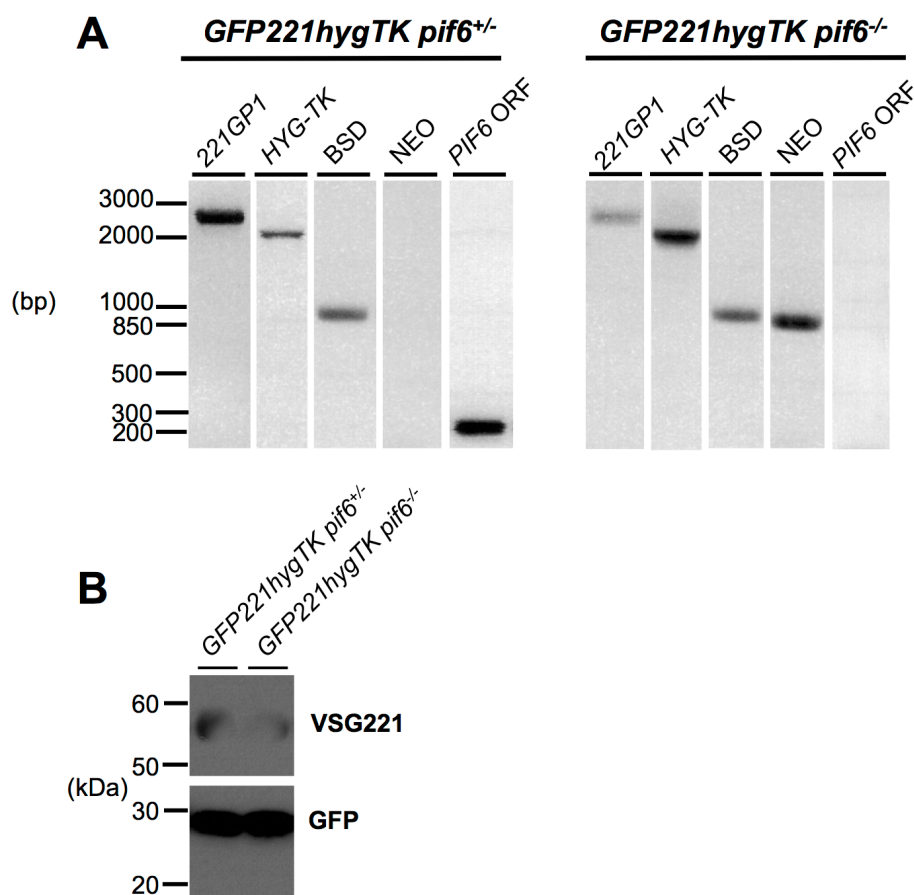
Repeated transformations to generate putative *GFP221hygTK mus81<sup>-/-</sup>* clones were attempted, which upon PCR analysis were shown to retain the *mus81* ORF. Multiple antibiotic drug concentrations in these transformations were used. Initial transformations used 10  $\mu\text{g.mL}^{-1}$  hygromycin, 5  $\mu\text{g.mL}^{-1}$  blasticidin and 2.5  $\mu\text{g.mL}^{-1}$  G418. Puromycin was also used in subsequent transformations, at a concentration at 0.2  $\mu\text{g.mL}^{-1}$ , in addition to these antibiotics. Again, antibiotic resistant clones were selected using higher concentrations of blasticidin and G418 than previous transformations: 10  $\mu\text{g.mL}^{-1}$  hygromycin, 10  $\mu\text{g.mL}^{-1}$  blasticidin, 5  $\mu\text{g.mL}^{-1}$  G418 and 0.2  $\mu\text{g.mL}^{-1}$  puromycin. The three transformations produced 21 putative *GFP221hygTK mus81<sup>-/-</sup>* clones, nine of which were analysed by PCR and all were correct.



**Figure 4-29** PCR and western blot analysis of *GFP221hygTK mus81* mutants. Genomic DNA was extracted from *GFP221hygTK mus81<sup>+/-</sup>* and *GFP221hygTK mus81<sup>-/-</sup>* cells. A 214 bp region of the ORF (primers #79 & #80) was PCR amplified (“MUS81 ORF”) to test for presence of the *MUS81* ORF. PCR amplification using primers #52 & #154 (966 bp PCR product) that bind in the *BSD* resistance cassette and upstream of *MUS81* (“BSD”) was used to test for correct integration of the *BSD* knockout construct. PCR amplification using primers #52 and #155 (974 bp PCR product) that bind in the *NEO* resistance cassette and upstream of *MUS81* (“NEO”) was used to test for correct integration of the *NEO* knockout construct. Presence of the 221GP1 construct was tested by PCR amplification using primers #120 and #120 (2.8 kbp PCR product, “221GP1”) and presence of the HYG-TK construct was tested by PCR amplification using primers #131 and #133 (2.2 kbp PCR product, “HYG-TK”) constructs. Primer sequences can be found in Appendix 7.1. Gaps

indicate that lanes have been aligned in this figure after excision from multiple gels or from disparate parts of the same gel; size markers are shown (ladder, bp).

Multiple transformations of *GFP221hygTK pif6<sup>+/-</sup>* with  $\Delta PIF6::NEO$  were attempted in order to successfully obtain a *GFP221hygTK pif6<sup>-/-</sup>* clone. Initial transformation of *GFP221hygTK pif6<sup>+/-</sup>* with  $\Delta PIF6::NEO$  used the following antibiotic drug concentrations: 5  $\mu\text{g}.\text{mL}^{-1}$  hygromycin, 5  $\mu\text{g}.\text{mL}^{-1}$  blasticidin and 2.5  $\mu\text{g}.\text{mL}^{-1}$  G418. These concentrations were altered to 5  $\mu\text{g}.\text{mL}^{-1}$  hygromycin, 5  $\mu\text{g}.\text{mL}^{-1}$  blasticidin, 2.5  $\mu\text{g}.\text{mL}^{-1}$  G418 and 0.2  $\mu\text{g}.\text{mL}^{-1}$  puromycin when a further transformation was attempted. Putative *GFP221hygTK pif6<sup>-/-</sup>* clones obtained were shown by PCR analysis to retain the *PIF6* ORF (data not shown). Upon increasing the selective antibiotic concentrations in the transformation to 10  $\mu\text{g}.\text{mL}^{-1}$  hygromycin, 5  $\mu\text{g}.\text{mL}^{-1}$  blasticidin, 2.5  $\mu\text{g}.\text{mL}^{-1}$  G418 and 0.2  $\mu\text{g}.\text{mL}^{-1}$  puromycin one putative *GFP221hygTK pif6<sup>-/-</sup>* clone was obtained. As shown in Figure 4-30A, successful PCR amplification of the “NEO”, “BSD”, “HYG-TK” and “GFP” PCR products in the putative *GFP221hygTK pif6<sup>-/-</sup>* clone confirmed that the  $\Delta PIF6::NEO$  and  $\Delta PIF6::BSD$  cassettes were integrated correctly relative to *PIF6* and that the clone contained the HYG-TK and 221GP1 constructs. Furthermore, the *PIF6* ORF could no longer be PCR amplified, indicating successful knockout of *PIF6*. Additionally, western blot analysis (Fig. 4-30B) of VSG221 and *eGFP* expression confirmed that the *GFP221hygTK pif6<sup>-/-</sup>* clone still expressed VSG221 and *eGFP*.



**Figure 4-30** PCR and western blot analysis of *GFP221hygTK pif6* mutants (A) Genomic DNA was extracted from *GFP221hygTK*, *GFP221hygTK pif6<sup>+/-</sup>* and *GFP221hygTK pif6<sup>-/-</sup>* cells. A 245 bp region of the ORF (primers #81 & #82) was PCR amplified (“*PIF6* ORF”) to test for presence of the *RECQ2* ORF. PCR amplification using primers #53 & #154 (890 bp PCR product) that bind in the *BSD* resistance cassette and upstream of *RECQ2* (“*BSD*”) was used to test for correct integration of the *BSD* knockout construct. PCR amplification using primers #53 and #155 (898 bp PCR product) that bind in the *NEO* resistance cassette and upstream of *RECQ2* (“*NEO*”) was used to test for correct integration of the *NEO* knockout construct. Presence of the 221GP1 construct was tested by PCR amplification using primers #120 and #120 (2.8 kbp PCR product, “221GP1”) and presence of the *HYG-TK* construct was tested by PCR amplification using primers #131 and #133 (2.2 kbp PCR product, “*HYG-TK*”) constructs. Primer sequences can be found in Appendix 7.1. Size markers are shown (ladder, bp). (B) Total protein extracts were separated by SDS-PAGE and western blotted. Membranes were probed with rabbit anti-VSG221 antiserum (1:20,000 dilution, (Glover *et al.*, 2013a)) and rabbit anti-GFP antiserum (1:5000 dilution, abcam). Both membranes were probed with goat HRP-conjugated anti-rabbit antiserum (Life Technologies). Size markers (kDa) are shown. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels/membranes or from disparate parts of the same gel/membrane.

#### 4.7.4 VSG switching analysis

To analyse VSG switching in *recq2*, *mus81* and *pif6* mutants, VSG switching assays were performed as described in Materials and Methods 2.2.4. Cells were diluted to a density of  $1 \times 10^4$  cells.mL<sup>-1</sup>, either in the presence or absence of puromycin (0.2 µg.mL<sup>-1</sup>) but otherwise free of antibiotics, and cultured for 48 hours. Absence of hygromycin selection allowed cells to switch VSG.

Puromycin selection was used in some cultures in order to eliminate switching events that result in the deletion of the VSG BES, a phenomenon frequently observed in *in vitro* switching experiments, and to select against *in situ* switching (Benmerzouga *et al.*, 2013; Cross *et al.*, 1998; Jehi *et al.*, 2014b; Kim & Cross, 2010; Kim & Cross, 2011; Povelones *et al.*, 2012; Rudenko *et al.*, 1998). Following 48 hours of incubation, cultures were diluted to either  $5 \times 10^3$  cells.mL<sup>-1</sup> or  $2.5 \times 10^3$  cells.mL<sup>-1</sup> and ganciclovir (Sigma) was added to a final concentration of 4 µg.mL<sup>-1</sup>. The cell concentration used in each experiment is indicated in the appropriate sections below. Cultures were then distributed across three 96 well plates, each well containing 200 µL of culture.

After incubation for seven days, the number of surviving wells per plate was counted. The VSG switching frequency was calculated as described in Materials and Methods, Section 2.2.4.2. Briefly, the number of survivors (totaled from the three 96 well plates) was divided by the total number of cells plated to obtain the number of switching events per cell, as each surviving well is a proxy for one switching event. The number of switching events per cell was then divided by the number of population doublings that had occurred in the 48 hour incubation period giving the number of switching events/cell/generation. This was done to adjust the calculated VSG switching rate to account for the slower growth of some mutants in the GFP221hygTK background. The number of population doublings was calculated for each culture in every experiment using the cell densities at the beginning and end of the 48 hour incubation period.

A selection of surviving clones were then analysed to determine the different VSG switching mechanisms used, as described in Materials and Methods, Sections 2.2.4.2 and 2.5.2. VSG221 and eGFP expression were assayed by western blotting, using anti-VSG221 antiserum (gift, David Horn, (Glover *et al.*, 2013a)) and an anti-GFP antiserum (abcam). Ponceau staining of the nitrocellulose membrane was carried out to confirm that protein was present on the blot for clones that were VSG221<sup>+</sup> and eGFP<sup>+</sup>. The presence or absence of the VSG221 and eGFP genes was determined by PCR-amplification of a region of the ORF of each gene. Primers #156 and #157 were used to amplify a 522 bp region of VSG221 and primers #160 and #161 were used to amplify a 361 bp region of eGFP. A 453 bp region of the RNA polymerase I (RNA PolI) large subunit (Tb427.08.5090) was amplified by PCR, using primers #158 and #159, as a

positive control in all clones. *RNA Poll* and *VSG221* PCR primer sequences were those as used in Povelones *et al.* (2012). The results obtained from the *GFP* PCR were inconclusive due to the poor efficiency of the PCR, and it was unclear whether clones contained the *GFP* gene using this PCR. The *GFP* PCR data are therefore not discussed in the text, though the PCR gels are shown in Figure 4-32 and Figure 4-35 and the results, based on the best estimation of the presence of amplification products in these gels, are included in the summaries in Tables 4-2, 4-4 and 4-6.

#### 4.7.4.1 Analysis of VSG switching in *recq2* mutants

Two experiments were performed to investigate the effect of mutation of *RECQ2* on the rate of VSG switching and profile of switching events. Experiments were performed as described above (see Section 4.8.4). In the first experiment, cell lines were seeded in 96 well plates at a density of  $5 \times 10^3$  cells.mL<sup>-1</sup>. Table 4-1 shows the number of survivors (wells with live cells) on each of the three 96 well plates for each cell line. The total number of survivors from the three 96 well plates was then used to calculate the VSG switching rate for each cell line in the presence or absence of puromycin (Fig. 4-31A). In the absence of puromycin, the switching rate of wild type cells was  $1.03 \times 10^{-5}$  events per cell per generation (events/cell/generation). This switching rate is comparable with the switching rates for wild type cells found by other studies using a similar VSG switching assay strategy (Kim & Cross, 2010; Kim & Cross, 2011; Povelones *et al.*, 2012), which were in the range  $1.05 \times 10^{-5}$  -  $2.8 \times 10^{-5}$ . However, the rate found in this experiment was considerably higher than the rate of  $\sim 5 \times 10^{-6}$  found by Benmerzouga *et al.* (2013). *recq2*<sup>+/-</sup> and *recq2*<sup>-/-</sup> cells each had an elevated VSG switching rate of  $2.35 \times 10^{-5}$  and  $2.93 \times 10^{-5}$  events/cell/generation respectively. The survival rate of each cell line (the percentage of wells containing live cells at the end of the incubation period) is shown in Table 4-1.

In the presence of puromycin, the VSG switching rate would be expected to decrease compared to in the absence puromycin, since switching events involving the inactivation or loss of the *VSG221* BES or *in situ* switching are suppressed. This was the case for *recq2*<sup>+/-</sup> and *recq2*<sup>-/-</sup> mutants, the switching rates of which decreased by approximately three-fold and two-fold,

respectively. However, the switching rate of wild type cells remained unchanged at  $1.03 \times 10^{-5}$  events/cell/generation.

A small selection of surviving clones were analysed to determine the VSG switching profiles of wild type and *recq2* mutant cell lines. Three clones each from the +puromycin and -puromycin conditions for wild type and *recq2*<sup>+/-</sup> cells, and six each from the +puromycin and -puromycin conditions for *recq2*<sup>-/-</sup> cells were analysed. Figure 4-31B shows the switching profile for wild type, *recq2*<sup>+/-</sup> and *recq2*<sup>-/-</sup> cells. Figures 4-32 and 4-33 show the PCR and western blot analysis results for each clone analysed. Table 4-2 is a summary of the analysis of all clones examined. As observed in a previous study using this cell line (Povelones *et al.*, 2012), and in other *in vitro* switching assays that relied upon immune selection *in vivo* (Bell & McCulloch, 2003; Hartley & McCulloch, 2008; McCulloch & Barry, 1999; Proudfoot & McCulloch, 2005), switching events involving deletion of the VSG BES (ES GC/*in situ* + ES del) predominated (two out of three clones analysed) in wild type cells in the absence of puromycin (Fig. 4-31B). The third clone was determined to be a *TK* mutant and to have not switched VSG. A high proportion of *TK* mutant survivors has been observed previously using this assay strategy (Povelones *et al.*, 2012). *TK* mutants were absent from the switching profile of analysed *recq2*<sup>+/-</sup> clones, where instead two out of three clones were *in situ* switchers, and the remainder VSG BES GC/*in situ* + ES deletion events. However, the switching profile of *recq2*<sup>-/-</sup> mutants was clearly different to that of wild type and *recq2*<sup>+/-</sup> cells. Two thirds of analysed clones had undergone VSG GC and the remaining third had undergone telomere XO, neither of which was observed in wild type or *recq2*<sup>+/-</sup> mutants and both of which are recombination-mediated switching events.

In the presence of puromycin, the switching profiles of wild type and *recq2* mutant cells were more similar to one another, and were altered relative to the profile seen in the absence of puromycin, despite the lack of detectable switching rate change in wild type cells (Fig. 4-31B). In these conditions, VSG GC events accounted for two thirds of switching events, with *TK* mutants accounting for the remaining third. Such a change is consistent with puromycin imposing a constraint on the assay by removing any processes in which *PUR* expression is lost. In the *recq2*<sup>-/-</sup> mutants, much less change in switching was seen, with telomere XO and VSG GC events each comprising half of the reactions

analysed. Thus, addition of puromycin altered the VSG switching profile of wild type and *recq2*<sup>+/-</sup> cells, but did not alter the switching profile of *recq2*<sup>-/-</sup> mutants, presumably because *recq2*<sup>-/-</sup> mutants preferentially underwent recombination-mediated switching events in the vicinity of *VSG221*, even in the absence of selection (by puromycin) for them.

| Cell line                   | PURO | Number of survivors |         |         |       | % survival |         |         |      |      |
|-----------------------------|------|---------------------|---------|---------|-------|------------|---------|---------|------|------|
|                             |      | Plate 1             | Plate 2 | Plate 3 | Total | Plate 1    | Plate 2 | Plate 3 | Mean | SEM  |
| Wild type                   | -    | 34                  | 39      | 31      | 104   | 35.4       | 40.6    | 32.3    | 36.1 | 2.43 |
|                             | +    | 31                  | 30      | 28      | 89    | 32.3       | 31.3    | 29.2    | 30.9 | 0.92 |
| <i>recq2</i> <sup>+/-</sup> | -    | 72                  | 69      | 62      | 203   | 75.0       | 71.9    | 64.6    | 70.5 | 3.09 |
|                             | +    | 21                  | 27      | 20      | 68    | 21.9       | 28.1    | 20.8    | 23.6 | 2.28 |
| <i>recq2</i> <sup>-/-</sup> | -    | 77                  | 61      | 73      | 211   | 80.2       | 63.5    | 76.0    | 73.3 | 5.01 |
|                             | +    | 41                  | 27      | 38      | 106   | 42.7       | 28.1    | 39.6    | 36.8 | 4.43 |

Table 4-1 Survival data from first *RECQ2* VSG switching assay

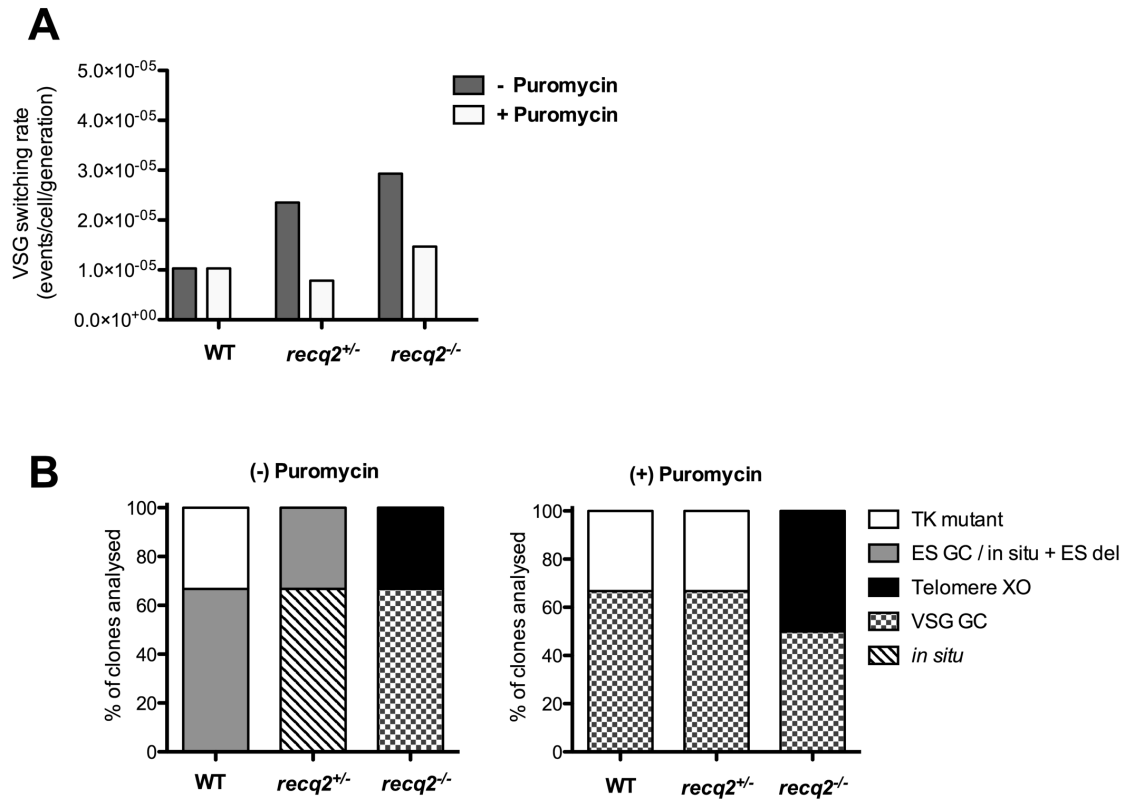
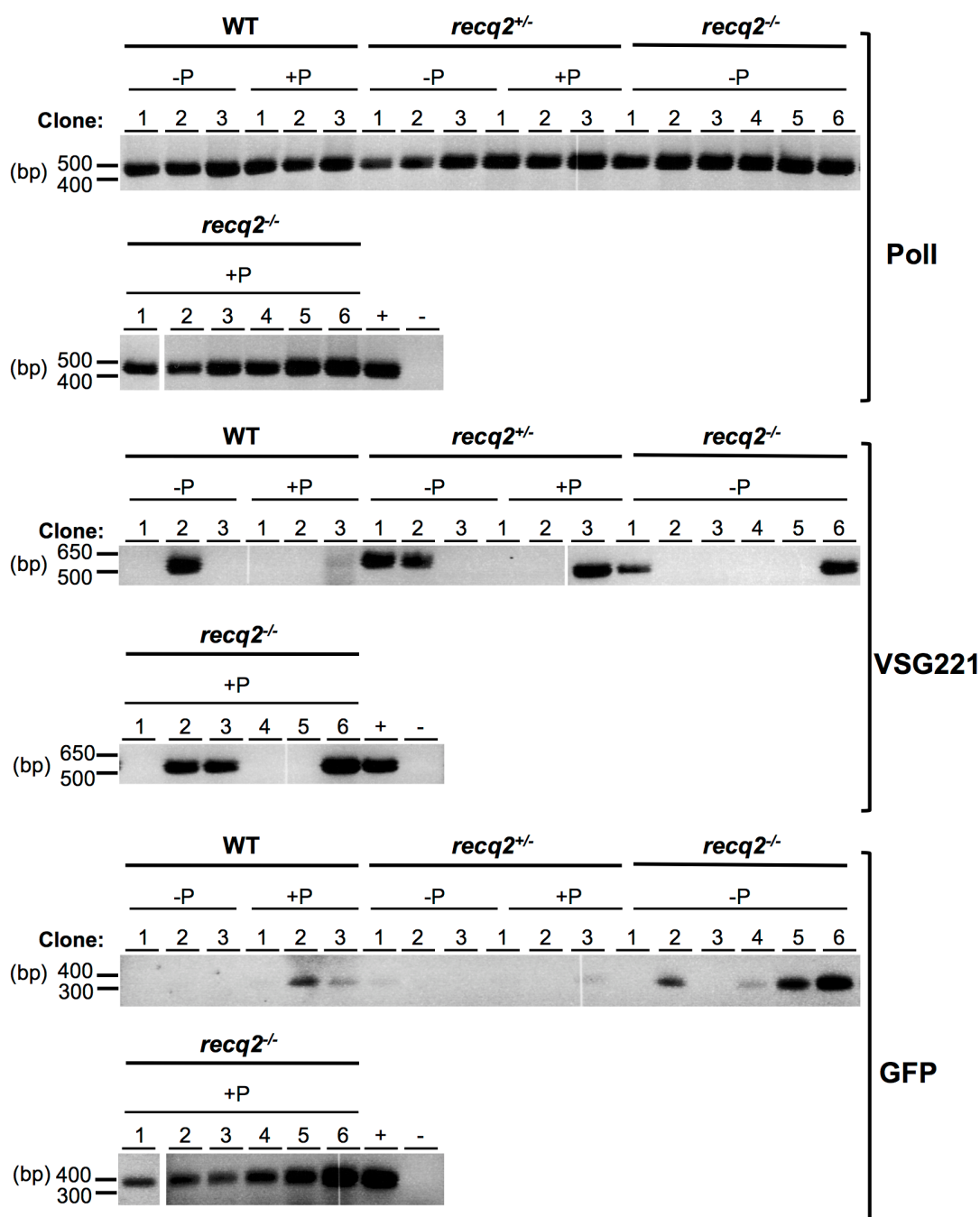


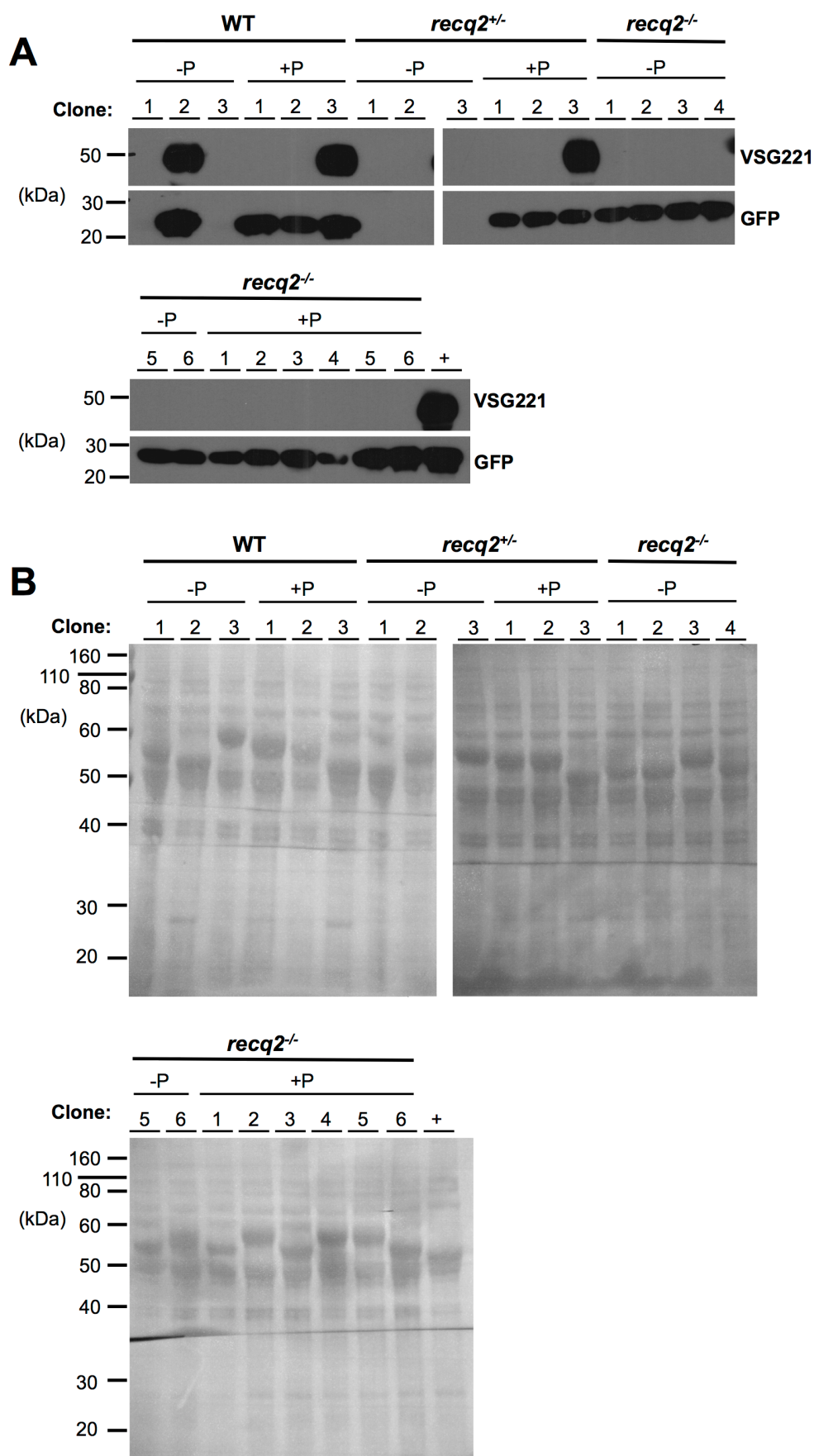
Figure 4-31 Switching rate and profile in first *RECQ2* VSG switching assay. (A) The switching rate of *GFP221hygTK* wild type and *GFP221hygTK recq2* mutants was calculated from the number of surviving clones as a proxy for VSG switching, following limiting dilution in the presence of ganciclovir and with (+) or without (-) puromycin. (B) The switching profiles of survivors in both the presence and absence of puromycin, represented as a percentage of total clones analysed. Switching events are as described in Figure 4-24.

**VSG gene conversion (VSG GC), telomere crossover (telomere XO), VSG BES gene conversion / *in situ* switch + ES deletion (ES GC/*in situ* + ES del)**



**Figure 4-32 PCR analysis of first *RECQ2* VSG switching experiment**  
 Genomic DNA was extracted from a selection of surviving GFP221hygTK wild type and GFP221hygTK *recq2* mutant clones. A 453 bp region of the RNA polymerase I (RNA Poll) large subunit (Tb427.08.5090) was amplified by PCR using primers #158 and #159 as a positive control for the presence of genomic DNA. Primers #156 and #157 were used to amplify a 522 bp region of VSG221 and primers #160 and #161 were used to amplify a 361 bp region of eGFP. GFP221hygTK wild type genomic DNA (+) was used as a positive control in all PCRs. Distilled water was used as a negative control (-). -P, clones recovered from the assay conducted without puromycin selection; +P, clones recovered from the assay conducted with puromycin selection. PCR primer sequences were those as used in (Povelones *et al.*, 2012) and can be found in Appendix 7.1. Size markers (bp, ladder) are shown. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels or from disparate parts of the same gel.





**Figure 4-33** Western analysis of first *RECQ2* VSG switching experiment  
**(A)** Western analysis of a selection of surviving *GFP221hygTK* wild type and *GFP221hygTK recq2* mutant clones. Total protein extracts were separated by SDS-PAGE and western

blotted. Membranes were probed with rabbit anti-VSG221 antiserum (1:20,000 dilution, (Glover *et al.*, 2013a)) and rabbit anti-GFP antiserum (1:5000 dilution, abcam). Both membranes were probed with goat HRP-conjugated anti-rabbit antiserum (Life Technologies). (-). -P, clones recovered from the assay conducted without puromycin selection; +P, clones recovered from the assay conducted with puromycin selection. GFP221hygTK wild type cells were used as a positive control (+). (B) Ponceau staining of the membranes shown in (A) after transfer. Membranes were incubated in ponceau solution for 10 minutes, rinsed and imaged. Size markers (kDa) are shown. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels/membranes or from disparate parts of the same gel/membrane.

| Cell line                   | Puromycin | Clone | Survival in: |     | PCR  |        |     | Western |     | Type of switching event       |
|-----------------------------|-----------|-------|--------------|-----|------|--------|-----|---------|-----|-------------------------------|
|                             |           |       | PUR          | HYG | POLI | VSG221 | GFP | VSG221  | GFP |                               |
| WT                          | -         | 1     |              |     |      |        |     |         |     | ES GC/ <i>in situ</i> +ES del |
|                             |           | 2     |              |     |      |        |     |         |     | TK mutant                     |
|                             |           | 3     |              |     |      |        |     |         |     | ES GC/ <i>in situ</i> +ES del |
|                             | +         | 1     |              |     |      |        |     |         |     | VSG GC                        |
|                             |           | 2     |              |     |      |        |     |         |     | VSG GC                        |
|                             |           | 3     |              |     |      |        |     |         |     | TK mutant                     |
| <i>recq2</i> <sup>+/-</sup> | -         | 1     |              |     |      |        |     |         |     | <i>in situ</i>                |
|                             |           | 2     |              |     |      |        |     |         |     | <i>in situ</i>                |
|                             |           | 3     |              |     |      |        |     |         |     | ES GC/ <i>in situ</i> +ES del |
|                             | +         | 1     |              |     |      |        |     |         |     | VSG GC                        |
|                             |           | 2     |              |     |      |        |     |         |     | VSG GC                        |
|                             |           | 3     |              |     |      |        |     |         |     | TK mutant                     |
| <i>recq2</i> <sup>-/-</sup> | -         | 1     |              |     |      |        |     |         |     | Telomere XO                   |
|                             |           | 2     |              |     |      |        |     |         |     | VSG GC                        |
|                             |           | 3     |              |     |      |        |     |         |     | VSG GC                        |
|                             |           | 4     |              |     |      |        |     |         |     | VSG GC                        |
|                             |           | 5     |              |     |      |        |     |         |     | VSG GC                        |
|                             |           | 6     |              |     |      |        |     |         |     | Telomere XO                   |
|                             | +         | 1     |              |     |      |        |     |         |     | VSG GC                        |
|                             |           | 2     |              |     |      |        |     |         |     | Telomere XO                   |
|                             |           | 3     |              |     |      |        |     |         |     | Telomere XO                   |
|                             |           | 4     |              |     |      |        |     |         |     | VSG GC                        |
|                             |           | 5     |              |     |      |        |     |         |     | VSG GC                        |
|                             |           | 6     |              |     |      |        |     |         |     | Telomere XO                   |

 Negative  
 Positive

**Table 4-2** Summary of the analysis of clones from first *RECQ2* VSG switching experiment

Summary of the data shown in Figures. 4-30 and 4-31. PUR, puromycin (1  $\mu\text{g}.\text{mL}^{-1}$ ); HYG, hygromycin (10  $\mu\text{g}.\text{mL}^{-1}$ ). Telomere XO, telomere crossover; VSG GC, VSG gene conversion. Blue indicates a negative result, yellow indicates a positive result.

The switching assay above was repeated as described, except that cells were diluted to a density of  $2.5 \times 10^3$  cells. $\text{mL}^{-1}$ . This was to attempt to ensure that *recq2*<sup>-/-</sup> survivors obtained were clonal, given the high number of positive wells in *recq2*<sup>-/-</sup> plates obtained from the higher cell density plating of  $5 \times 10^3$  cells. $\text{mL}^{-1}$  in the first experiment. Table 4-3 shows the number of survivors (wells with live cells) for each cell line. Additionally, as a control to ensure that the ganciclovir killed TK-positive cells, GFP221hygTK wild type, *recq2*<sup>+/-</sup> and *recq2*<sup>-/-</sup> cells that had been continuously cultured in the presence of hygromycin and puromycin were diluted and treated with ganciclovir in the same manner as switching assay cultures. Following 10 days incubation there were no surviving wells (data not shown).

| Cell line                   | PURO | Number of survivors |         |         |       | % relative survival |         |         |      |      |
|-----------------------------|------|---------------------|---------|---------|-------|---------------------|---------|---------|------|------|
|                             |      | Plate 1             | Plate 2 | Plate 3 | Total | Plate 1             | Plate 2 | Plate 3 | Mean | SEM  |
| Wild type                   | -    | 26                  | 19      | 24      | 69    | 27.1                | 19.8    | 25.0    | 24.0 | 2.17 |
|                             | +    | 9                   | 5       | 8       | 22    | 9.4                 | 5.2     | 8.3     | 7.6  | 1.25 |
| <i>recq2</i> <sup>+/-</sup> | -    | 10                  | 16      | 18      | 44    | 10.4                | 16.7    | 18.8    | 15.3 | 2.50 |
|                             | +    | 5                   | 6       | 8       | 19    | 5.2                 | 6.3     | 8.3     | 6.6  | 0.92 |
| <i>recq2</i> <sup>-/-</sup> | -    | 41                  | 26      | 35      | 102   | 42.7                | 27.1    | 36.5    | 35.4 | 4.54 |
|                             | +    | 31                  | 44      | 45      | 120   | 32.3                | 45.8    | 46.9    | 41.7 | 4.70 |

**Table 4-3 Survival in a second *RECQ2* VSG switching assay**

The VSG switching rate of wild type cells in the absence of puromycin was  $1.92 \times 10^{-5}$  events/cell/generation (Fig. 4-34A), similar to the switching rate from the previous experiment ( $1.03 \times 10^{-5}$ ). The switching rate of *recq2*<sup>+/-</sup> and *recq2*<sup>-/-</sup> mutants was  $1.02 \times 10^{-5}$  events/cell/generation and  $3.54 \times 10^{-5}$  events/cell/generation respectively. Thus, only the switching rate of the +/- mutants differed substantially between the two experiments. Whether this change simply reflects fluctuation, given the low absolute number of switching events, is unclear.

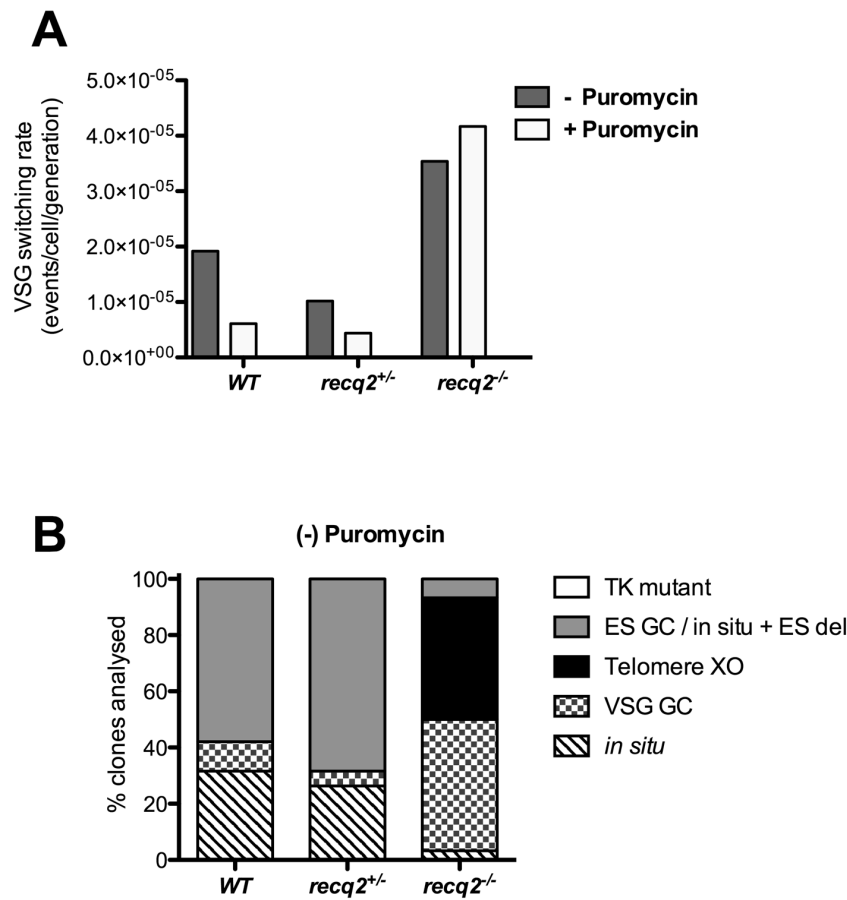
In the presence of puromycin, the switching rate of wild type and *recq2*<sup>+/-</sup> cells was decreased approximately three-fold and two-fold, respectively (Fig. 4-34A). This is the expected result as, (1) puromycin suppresses *in situ* and VSG BES deletion switching events, and (2) the switching profiles of wild type and *recq2*<sup>+/-</sup> cells in the first experiment were dominated by these types of switching events. However, this is different to the first experiment where the switching rate of wild type cells did not alter in the presence of puromycin. Again, this may simply reflect fluctuation. The switching rate of *recq2*<sup>-/-</sup> cells was essentially equivalent in the absence or presence of puromycin, increasing very slightly from  $3.54 \times 10^{-5}$  in the former conditions to  $4.17 \times 10^{-5}$  events/cell/generation in the latter (Fig. 4-34A). This result is once again different to the first experiment, where the switching rate of *recq2*<sup>-/-</sup> mutants decreased two-fold in the presence of puromycin. However, no change in the switching rate of *recq2*<sup>-/-</sup> cells in the presence of puromycin makes sense in the context of the switch profile of these cells in the first experiment, which

suggested a predominance of VSG-focussed recombination-based switches, rather than *in situ* or VSG BES deletion switches. Irrespective, the rate of switching in the absence of puromycin was very comparable between the two experiments.

A greater number of surviving clones were analysed to investigate the switching profiles than in the first experiment; 19 wild type, 19 *recq2*<sup>+/-</sup> and 30 *recq2*<sup>-/-</sup> clones were analysed, all from the non-puromycin condition. No clones were analysed from the puromycin treated assays because the addition of puromycin did not alter the switching profile of *recq2*<sup>-/-</sup> mutants in the first experiment. The VSG switching profiles of all three cell lines is shown in Figure 4-34B. In contrast to the first experiment, no *TK* mutants were recovered. As in the first experiment, the VSG switching profile of wild type cells was predominated (57.9%) by switches involving deletion of the active VSG BES (ES GC/*in situ* + ES del) (Fig. 4-34B). Analysis of a greater number of clones and the absence of *TK* mutants however, gives an expanded picture of the VSG switching profile. Approximately one third (31.6%) of wild type survivors had switched VSG by an *in situ* switch and the remaining 10.5% had switched VSG using VSG GC. The VSG switching profile of *recq2*<sup>+/-</sup> mutants was very similar to wild type cells, with the majority of survivors (68.4%) undergoing VSG BES deletion events and similar proportions of survivors switching VSG by VSG GC (5.3%) and *in situ* (26.3%) switches. Thus, the difference in switching profiles between these cells seen in the first experiment is most likely a consequence of the small number of clones examined.

The VSG switching profile of *recq2*<sup>-/-</sup> cells was, as in the first experiment, markedly different to wild type and *recq2*<sup>+/-</sup> cells. The switching profile was characterised by a dramatic increase in VSG switching events involving recombination, and a decrease in VSG BES deletion and *in situ* switching events. Approximately equal proportions of survivors underwent VSG GC (46.7%) and telomere XO (43.3%) events. This proportion of VSG GC is a four-fold increase from the proportion in wild type cells and an almost nine-fold increase from the proportion in *recq2*<sup>+/-</sup> cells, whilst telomere crossovers were not observed at all in wild type or *recq2*<sup>+/-</sup> survivors. The proportion of *recq2*<sup>-/-</sup> survivors that underwent *in situ* switching was 3.3% - nine-fold lower than that of wild type. Similarly, the proportion of *recq2*<sup>-/-</sup> survivors having undergone events involving

VSG BES deletion (ES GC/*in situ* + ES del) was 6.7%, an eight-fold reduction compared to wild type.



**Figure 4-34** Switching rate and profile of *recq2* mutants in a second VSG switching experiment  
(A) The switching rate of GFP221hygTK wild type and GFP221hygTK *recq2* mutants was calculated from the number of surviving clones, as a proxy for VSG switching, following treatment with ganciclovir after culture with or without puromycin. (B) The switching profiles of survivors of the different cell lines in both the puromycin and the non-puromycin condition, represented as a percentage of total surviving clones analysed.

The results of PCR analysis of each clone analysed in this experiment are shown in Figure 4-35 and the western analysis results are shown in Figure 4-36. In addition, a summary of the data in Figs. 4-35 and 4-36, as well the growth of each clone in 10 µg.mL<sup>-1</sup> hygromycin and 1 µg.mL<sup>-1</sup> puromycin and VSG switching mechanism inferred from these analyses is shown in Table 4-4.

**Figure 4-35** PCR analysis of second *RECQ2* VSG switching experiment

Genomic DNA was extracted from a selection of surviving GFP221hygTK wild type and *GFP221hygTK recq2* mutant clones. A 453 bp region of the RNA polymerase I (RNA PolI) large subunit (Tb427.08.5090) was amplified by PCR, using primers #158 and #159, as a positive control for the presence of genomic DNA. Primers #156 and #157 were used to amplify a 522 bp region of *VSG221* and primers #160 and #161 were used to amplify a 361 bp region of *eGFP*. GFP221hygTK wild type genomic DNA (+) was used as a positive control in all PCRs. Distilled water was used as a negative control (-). -P, clones recovered from the assay conducted without puromycin selection; +P, clones recovered from the assay conducted with puromycin selection. PCR primer sequences were those as used in (Povelones *et al.*, 2012) and can be found in Appendix 7.1. Size markers (bp, ladder) are shown. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels or from disparate parts of the same gel.

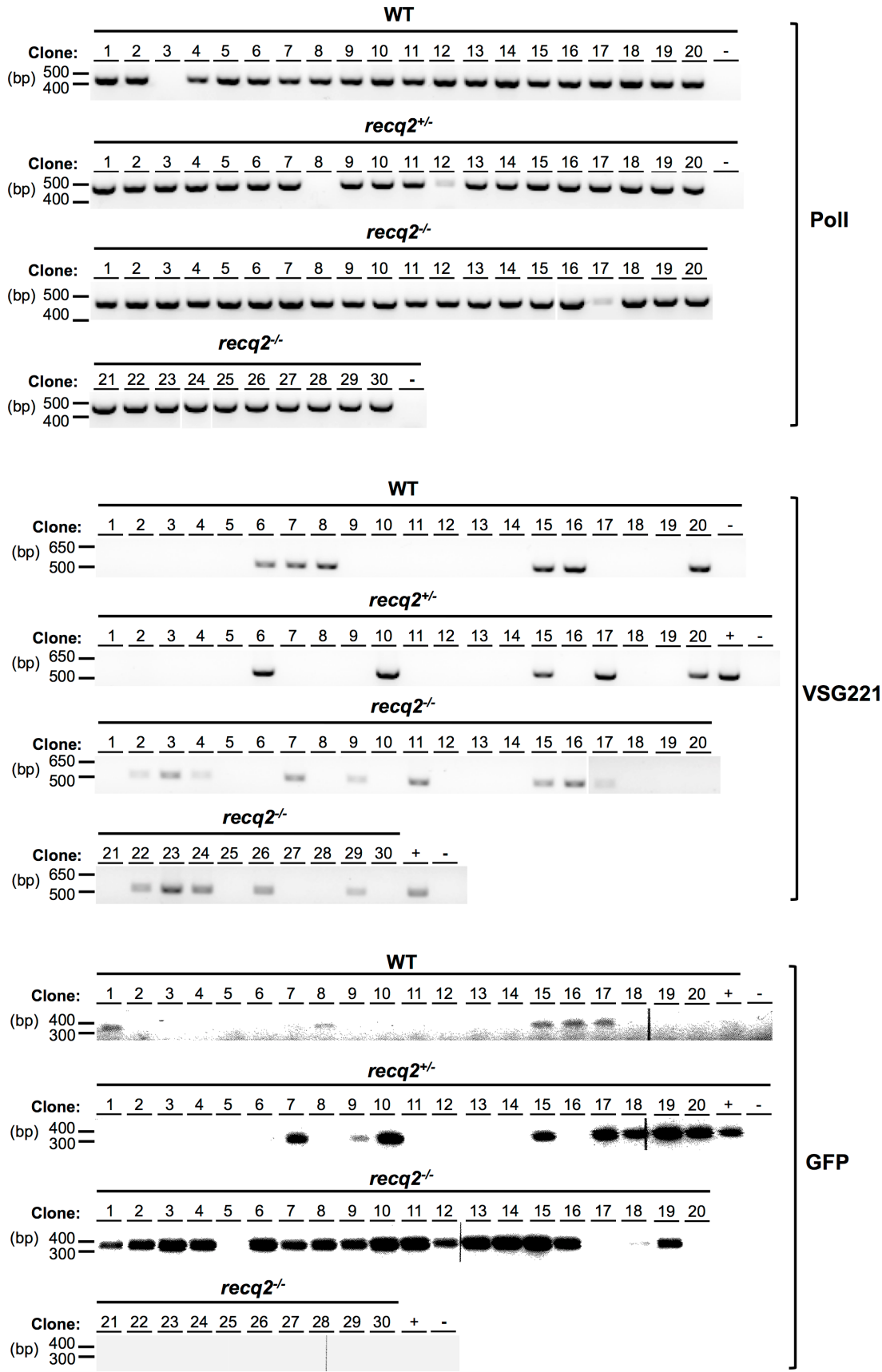


Figure 4-35

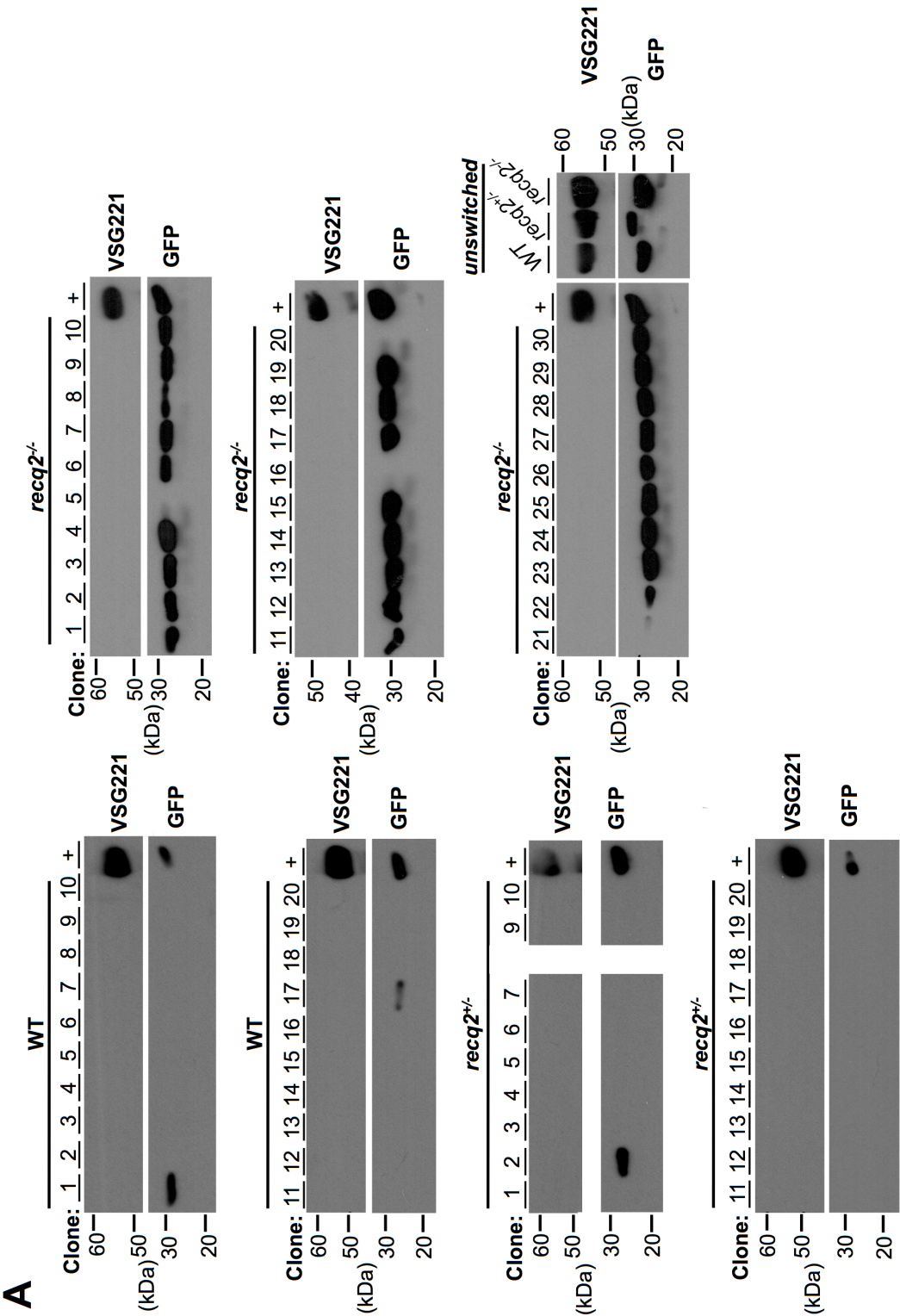


Figure 4-36 Western analysis of a second RECQ2 VSG switching experiment



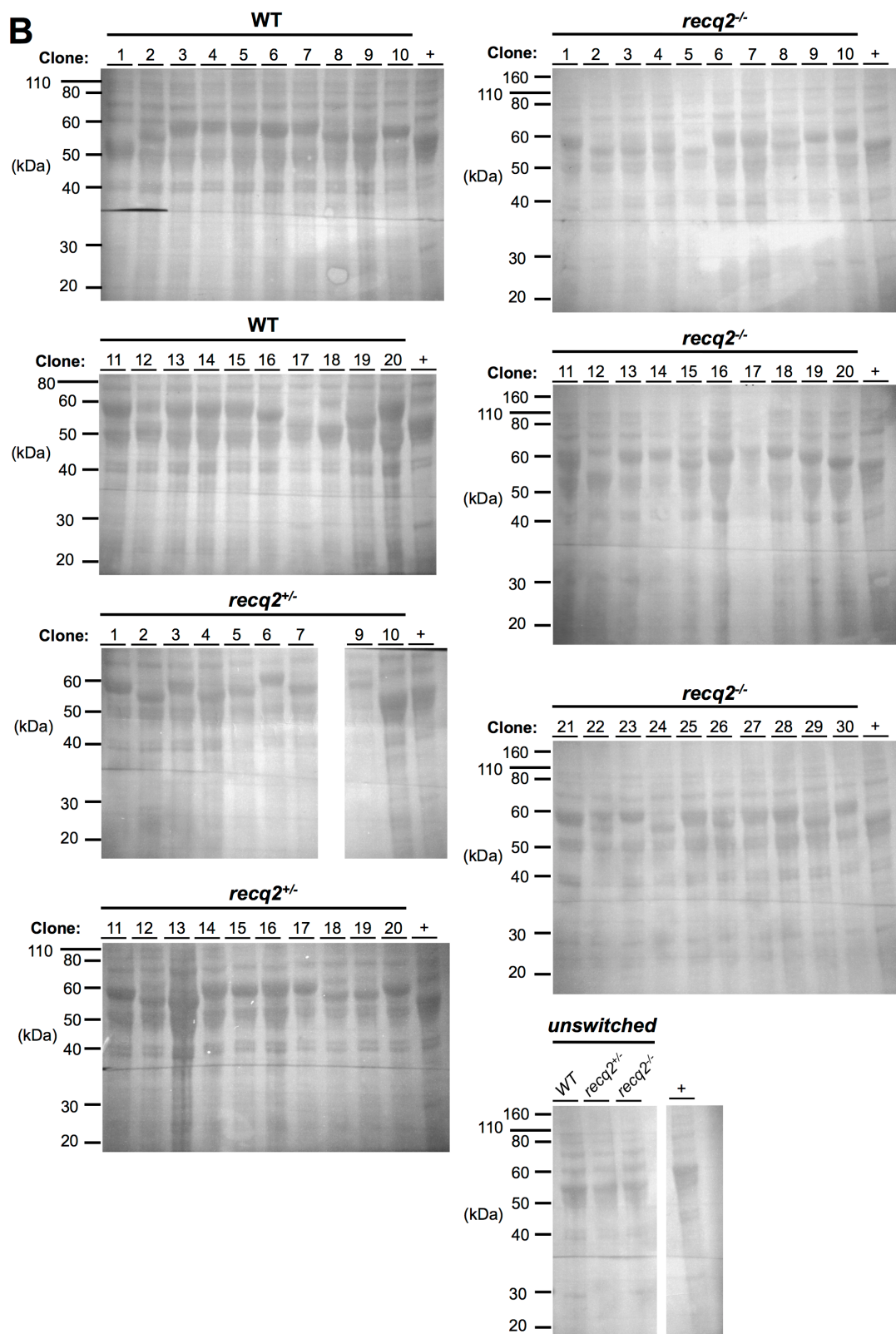




Figure 4-36 continued

**A)** Western analysis of a selection of surviving GFP221hygTK wild type and *GFP221hygTK recq2* mutant clones. Total protein extracts were separated by SDS-PAGE and western blotted. Membranes were probed with rabbit anti-VSG221 antiserum (1:20,000 dilution, (Glover *et al.*, 2013a)) and rabbit anti-GFP antiserum (1:5000 dilution, abcam). Both membranes were probed with goat HRP-conjugated anti-rabbit antiserum (Life Technologies). -P, clones recovered from the assay conducted without puromycin selection; +P, clones recovered from the assay conducted with puromycin selection.

**GFP221hygTK wild type cells were used as a positive control (+). (B) Ponceau staining of the membranes shown in (A) after transfer. Membranes were incubated in ponceau solution for 10 minutes, washed and imaged. Size markers (kDa) are shown. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels/membranes or from disparate parts of the same gel/membrane.**

| Cell line | Puromycin | Clone | PCR  |        |     | Western |     | Switching type          |
|-----------|-----------|-------|------|--------|-----|---------|-----|-------------------------|
|           |           |       | POLI | VSG221 | GFP | VSG221  | GFP |                         |
| WT        | -         | 1     |      |        |     |         |     | VSG GC                  |
|           |           | 2     |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 3     |      |        |     |         |     | not possible to analyse |
|           |           | 4     |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 5     |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 6     |      |        |     |         |     | <i>in situ</i>          |
|           |           | 7     |      |        |     |         |     | <i>in situ</i>          |
|           |           | 8     |      |        |     |         |     | <i>in situ</i>          |
|           |           | 9     |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 10    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 11    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 12    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 13    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 14    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 15    |      |        |     |         |     | <i>in situ</i>          |
|           |           | 16    |      |        |     |         |     | <i>in situ</i>          |
|           |           | 17    |      |        |     |         |     | VSG GC                  |
|           |           | 18    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 19    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 20    |      |        |     |         |     | <i>in situ</i>          |
| recq2+/-  | -         | 1     |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 2     |      |        |     |         |     | VSG GC                  |
|           |           | 3     |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 4     |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 5     |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 6     |      |        |     |         |     | <i>in situ</i>          |
|           |           | 7     |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 8     |      |        |     |         |     | not possible to analyse |
|           |           | 9     |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 10    |      |        |     |         |     | <i>in situ</i>          |
|           |           | 11    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 12    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 13    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 14    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 15    |      |        |     |         |     | <i>in situ</i>          |
|           |           | 16    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 17    |      |        |     |         |     | <i>in situ</i>          |
|           |           | 18    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 19    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 20    |      |        |     |         |     | <i>in situ</i>          |
| recq2-/-  | -         | 1     |      |        |     |         |     | VSG GC                  |
|           |           | 2     |      |        |     |         |     | Telomere XO             |
|           |           | 3     |      |        |     |         |     | Telomere XO             |
|           |           | 4     |      |        |     |         |     | Telomere XO             |
|           |           | 5     |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 6     |      |        |     |         |     | VSG GC                  |
|           |           | 7     |      |        |     |         |     | Telomere XO             |
|           |           | 8     |      |        |     |         |     | VSG GC                  |
|           |           | 9     |      |        |     |         |     | Telomere XO             |
|           |           | 10    |      |        |     |         |     | VSG GC                  |
|           |           | 11    |      |        |     |         |     | Telomere XO             |
|           |           | 12    |      |        |     |         |     | VSG GC                  |
|           |           | 13    |      |        |     |         |     | VSG GC                  |
|           |           | 14    |      |        |     |         |     | VSG GC                  |
|           |           | 15    |      |        |     |         |     | Telomere XO             |
|           |           | 16    |      |        |     |         |     | <i>in situ</i>          |
|           |           | 17    |      |        |     |         |     | Telomere XO             |
|           |           | 18    |      |        |     |         |     | VSG GC                  |
|           |           | 19    |      |        |     |         |     | VSG GC                  |
|           |           | 20    |      |        |     |         |     | ES GC/in situ + ES del  |
|           |           | 21    |      |        |     |         |     | VSG GC                  |
|           |           | 22    |      |        |     |         |     | Telomere XO             |
|           |           | 23    |      |        |     |         |     | Telomere XO             |
|           |           | 24    |      |        |     |         |     | Telomere XO             |
|           |           | 25    |      |        |     |         |     | VSG GC                  |
|           |           | 26    |      |        |     |         |     | Telomere XO             |
|           |           | 27    |      |        |     |         |     | VSG GC                  |
|           |           | 28    |      |        |     |         |     | VSG GC                  |
|           |           | 29    |      |        |     |         |     | Telomere XO             |
|           |           | 30    |      |        |     |         |     | VSG GC                  |

 Negative  
 Positive

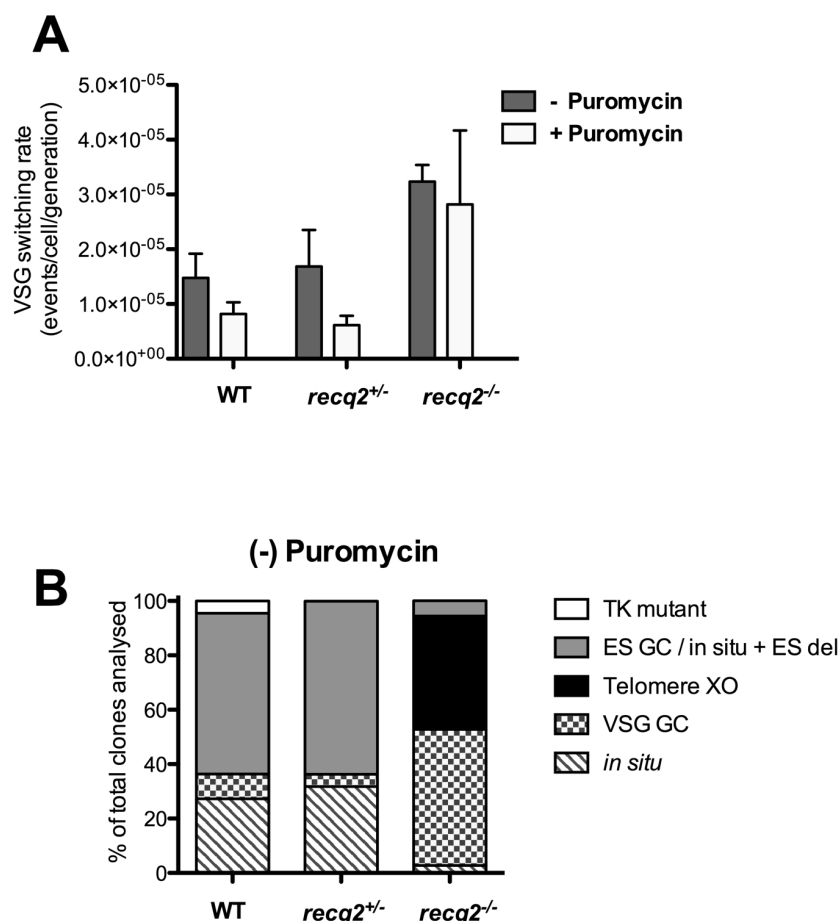
**Table 4-4** Summary of analysis of clones from a second *RECQ2* VSG switching experiment

Summary of the data shown in Figs. 4-36 and 4-37, and growth in puromycin ( $1 \mu\text{g.mL}^{-1}$ ) and hygromycin ( $10 \mu\text{g.mL}^{-1}$ ). *PUR*, puromycin ( $1 \mu\text{g.mL}^{-1}$ ); *HYG*, hygromycin ( $10 \mu\text{g.mL}^{-1}$ ).

**Telomere XO, telomere crossover; VSG GC, VSG gene conversion. Blue indicates a negative result, yellow indicates a positive result. Completely white rows indicate clones that could not be analysed, either due to absence of genomic DNA in genomic DNA extractions or cell lysates in western blot analysis.**

The two repeats of this assay are broadly similar. The VSG switching rates obtained for wild type cells, *recq2*<sup>+/-</sup> and *recq2*<sup>-/-</sup> in the absence of puromycin varied only a small amount between the two experiments. In both experiments it was observed that some surviving clones grew more slowly than others. This has been observed in other studies (Kim & Cross, 2010) but is currently not understood, though it has been suggested that slower growing clones may be switchers expressing ESAGs less advantageous for *in vitro* growth (Bitter *et al.*, 1998).

Due to the overall similarity between the experiments, the two datasets were combined (Fig. 4-37), allowing variation to be taken into consideration. Figure 4-38A shows that the VSG switching rate of *recq2*<sup>-/-</sup> cells was higher than that of wild type by approximately two-fold and was approximately similarly higher than *recq2*<sup>+/-</sup> mutants. The VSG switching rate of the *recq2*<sup>+/-</sup> cells showed greater variation between the two experiments and perhaps showed a trend towards a slight increase relative to wild type (which may be consistent with the minor growth impairment and MMS sensitivity phenotypes in these cells; Chapter 3), but this was not statistically significant. In the presence of puromycin, the combined data show that the rate of VSG switching was decreased by 1.8- and 2.7-fold in wild type and *recq2*<sup>+/-</sup> lines, respectively. In contrast, the same pronounced reduction was not seen in the VSG switching rate of *recq2*<sup>-/-</sup> mutants in the presence of puromycin, though the variation in these data between the two experimental repeats makes it difficult to draw firm conclusions. However, the high proportion of recombination-based switching events, which are not selected against in the puromycin-free switching profile of *recq2*<sup>-/-</sup> cells (discussed below), suggests that the second data set is accurate and that puromycin does not affect the VSG switching rate in *recq2*<sup>-/-</sup> cells.



**Figure 4-37 Switching rate and switching profile of *recq2* mutants from a combined dataset**

(A) The mean switching rate of GFP221hygTK wild type and GFP221hygTK *recq2* mutants was calculated from the mean number of surviving from the two experiments, following treatment with ganciclovir after culture with or without puromycin. Error bars represent standard error of the mean (SEM). (B) The switching profiles of survivors of the different cell lines in the non-puromycin condition, represented as a percentage of total surviving clones analysed from the two datasets.

The increase in VSG switching rate and VSG GC and crossover events in *recq2*<sup>-/-</sup> mutants is very similar to the phenotype described for *topo3a* and *rmi1* mutants (Kim & Cross, 2010; Kim & Cross, 2011). The increase in switching frequency observed in *recq2*<sup>-/-</sup> mutants (two to three-fold) was not as high as the elevated switching frequency reported in *rmi1* and *topo3a* null mutants (four-fold and 10 to 40-fold, respectively). However, the proportion of VSG GC and telomere XO switchers in *recq2*<sup>-/-</sup> mutants (~50% and ~40% respectively) is similar to those of *rmi1* (70% and 25%, respectively) and *topo3a* (70% and 23%, respectively) mutants, as well as all of these mutants sharing an almost complete absence of events leading to VSG BES loss (Kim & Cross, 2010; Kim & Cross, 2011). Although the switching rate of wild type cells in the Kim & Cross studies (2010; 2011) was broadly similar to the data presented here, the switching profile of wild type

cells was somewhat different. For example, *in situ* switchers were almost entirely absent (<2% of total switchers) in the Kim & Cross studies (Kim & Cross, 2010; Kim & Cross, 2011), whereas they constituted ~27% of wild type switchers in this work. Conversely, ~65% of wild type switchers recovered in Kim & Cross (2010) used VSG GC, as opposed to only ~10% here. These differences may be a result of differences in the cell lines. Although the strategies used are the same, the constructs used to generate the VSG switching lines were different. Small differences in the constructs, their expression levels and/or their integration into the active VSG BES may influence the VSG switching rate and switching profile. Nonetheless, the similarity in the elevation of VSG-focused recombination in the switching profiles of *recq2*, *topo3a* and *rmi1* null mutants is notable, given that has been suggested that a RECQ protein functions in a complex with RMI1 and TOPO3 $\alpha$ , forming the *T. brucei* homologue of the RTR complex. The RTR complex is composed of a RecQ helicase (BLM in mammals and SGS1 in *S. cerevisiae*, topoisomerase III  $\alpha$  (Topo3 $\alpha$ ) and RMI1/2 (BLAP75/18 in mammals) and resolves crossover intermediates such as HJs (Hickson & Mankouri, 2011; Singh *et al.*, 2008). Kim & Cross (2011) demonstrated that TbTOPO3 $\alpha$  (Tb927.8.3810) and TbRMI1 (Tb927.3.1830) interact with one another, but did not reveal interaction with any helicase. As RECQ1 is essential and may be involved in nuclear DNA replication (discussed in Chapter 5), the pronounced similarity of the phenotypes seen here suggest it is likely that RECQ2 is the more probable candidate to interact in a putative *T. brucei* RTR complex. Confirmation of this hypothesis would require co-immunoprecipitation of RECQ2 and RMI1/TOPO3 $\alpha$  and analysis of the switching rate and phenotype of double mutants.

#### 4.7.4.2 Analysis of VSG switching in *pif6* mutants

VSG switching analysis of *pif6* mutants was performed by the same approach as used in the second *recq2* mutant VSG switching assay (see Section 4.7.4.1); cells were seeded in the presence of ganciclovir (4  $\mu\text{g.mL}^{-1}$ ) at a density of  $2.5 \times 10^3$  cells.mL $^{-1}$  and 20 wild type, 20 *pif6* $^{+/-}$  and 30 *pif6* $^{-/-}$  surviving clones were analysed.

Table 4-5 shows the number of survivors on each of the three 96 well plates for wild type, *pif6* $^{+/-}$  and *pif6* $^{-/-}$  cell lines used to calculate the VSG switching rate

for each cell line (Fig. 4-38A). The VSG switching rate of wild type cells was measured at  $1.69 \times 10^{-5}$  events/cell/generation, similar to the switching rate we found for wild type cells previously ( $1.03 \times 10^{-5}$  -  $1.92 \times 10^{-5}$ , see Section 4.8.4.1). Deletion of one *PIF6* allele did not alter the VSG switching frequency compared to wild type, which was measured at  $1.55 \times 10^{-5}$  switches/cell/generation in *pif6*<sup>+/-</sup> cells. However, deletion of the second *PIF6* allele in *pif6*<sup>-/-</sup> cells increased the VSG switching rate two-fold to  $3.17 \times 10^{-5}$  switches/cell/generation, a similar elevation in VSG switching as was seen for *recq2*<sup>-/-</sup> mutants. In the presence of puromycin, the switching rate decreased approximately four-fold in wild type cells ( $3.89 \times 10^{-6}$ ) and approximately three-fold in *pif6*<sup>+/-</sup> cells ( $4.63 \times 10^{-6}$ ), as expected if VSG BES GC/deletion and *in situ* switching events predominate in these cells and were suppressed (see below). By contrast, the switching rate of *pif6*<sup>-/-</sup> did not change ( $3.43 \times 10^{-5}$ ) in the presence of puromycin.

| Cell line                  | PURO | Number of survivors |         |         |       | % relative survival |         |         |      |      |
|----------------------------|------|---------------------|---------|---------|-------|---------------------|---------|---------|------|------|
|                            |      | Plate 1             | Plate 2 | Plate 3 | Total | Plate 1             | Plate 2 | Plate 3 | Mean | SEM  |
| Wild type                  | -    | 26                  | 18      | 17      | 61    | 21.1                | 18.8    | 17.7    | 21.2 | 2.97 |
|                            | +    | 6                   | 6       | 2       | 14    | 6.3                 | 6.3     | 2.1     | 4.9  | 1.39 |
| <i>pif6</i> <sup>+/-</sup> | -    | 17                  | 28      | 22      | 67    | 17.7                | 29.2    | 22.9    | 23.3 | 3.31 |
|                            | +    | 6                   | 8       | 6       | 20    | 6.3                 | 8.3     | 6.3     | 6.9  | 0.69 |
| <i>pif6</i> <sup>-/-</sup> | -    | 47                  | 37      | 53      | 137   | 49.0                | 38.5    | 55.2    | 47.6 | 4.86 |
|                            | +    | 50                  | 48      | 50      | 148   | 52.1                | 50.0    | 52.1    | 51.4 | 0.69 |

**Table 4-5** Number of survivors and survival rate in PIF6 VSG switching assay

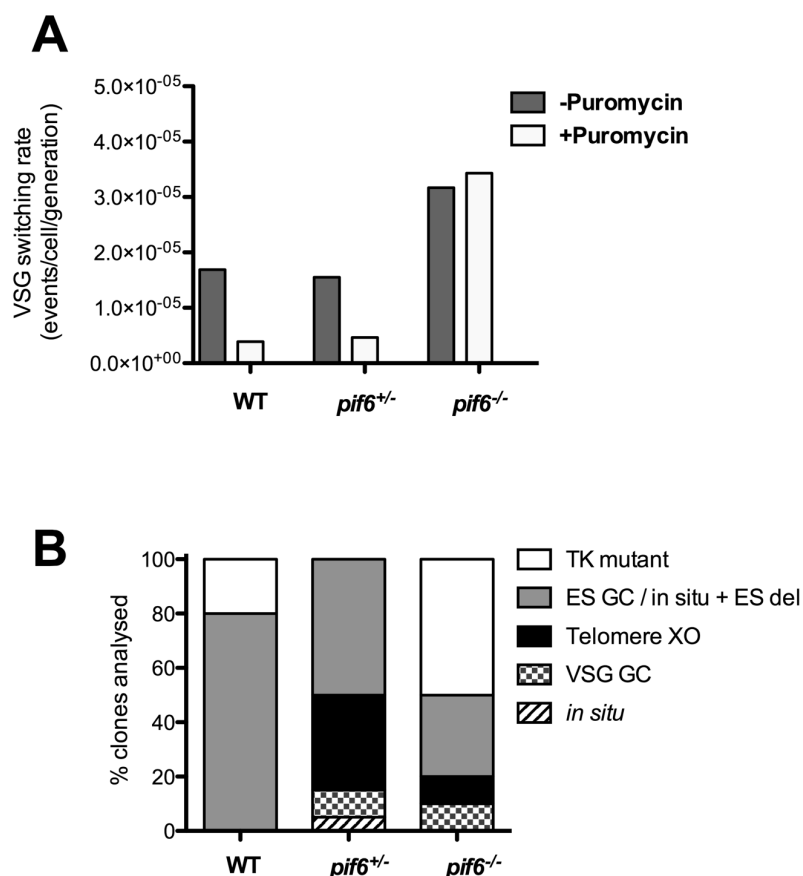
Analysis of survivors was undertaken similarly to that of the second RECQ2 VSG switching experiment described above (Section 4.7.4.1), except that *GFP* PCR analysis was not performed due to difficulties with the PCR. Clones were analysed by PCR for the presence of *RNA PolII* and *VSG221* and by western blot analysis for *GFP* and *VSG221* expression (Fig. 4-38B). Twenty wild type, 20 *pif6*<sup>+/-</sup> survivors and 30 *pif6*<sup>-/-</sup> survivors from the non-puromycin condition were analysed. PCRs and western blots of these data are shown in Figures 4-39 and 4-40 and a summary of the results for each clone is shown in Table 4-6.

In contrast to the previous VSG switching experiments described in Section 4.7.4.1, no VSG switching events other than VSG BES loss and *TK* mutants were detected in wild type cells. This may reflect inherent variability in the assay. One clone classified as a *TK* mutant was negative for *VSG221* by PCR but this probably reflects a failure in the PCR reaction as the clone expressed *VSG221* and *GFP* by western blot analysis. Loss of one *PIF6* allele did not alter the VSG switching rate however, the profile of *pif6*<sup>+/-</sup> mutants was different to that of wild type cells. Though the frequency of VSG BES loss events was high (50%), the remainder of VSG switching events detected were primarily telomere crossover and VSG GC (35% and 10% respectively), with a single *in situ* switching event detected (5%). This switching profile greatly resembles that observed in wild type cells in the second RecQ2 VSG switching experiment in Section 4.7.4.1 (Figure 4-34). The VSG switching profile of *pif6*<sup>-/-</sup> cells was dominated by *TK* mutants (50%). The majority of analysed clones that had switched VSG however had switched via a VSG BES loss event (30%), with the remainder of clones switching by VSG GC or telomere crossover (10% each).

Analysis of these switching data for *pif6* mutants is challenging for several reasons. Though in this experiment the VSG switching profile of wild type and *pif6*<sup>+/-</sup> cells is quite different, and suggests an increase in recombination reactions near the active VSG, the profile of *pif6*<sup>+/-</sup> cells actually closely resembles that of wild type cells observed previously (Fig 4-34). Secondly, the VSG switching profile of *pif6*<sup>-/-</sup> looks broadly similar to that of *pif6*<sup>+/-</sup> cells when the *TK* mutants are excluded. The high frequency of *TK* mutants in *pif6*<sup>-/-</sup> survivors is also interesting. These clones were classified as *TK* mutants because they retained the *VSG221* gene and maintained expression of *VSG221* and *GFP*. However, the high frequency of *pif6*<sup>-/-</sup> survivors here is reminiscent of the retention of *VSG221* observed in *HRES pif6*<sup>-/-</sup> mutants (Section 4.5.3, Fig. 4-15). Presumably the only way this could occur is if DSB induction in HRES and the initiating lesion in VSG switching both get repaired by an end-joining mechanism (such as MMEJ) in *pif6*<sup>-/-</sup> cells, rather than by recombination. However, given the relatively high percentage of *TK* mutants recovered in wild type cells in this experiment (20%), the 50% of *TK* mutants recovered in *pif6*<sup>-/-</sup> cells may simply reflect variability in *TK* mutant generation in this assay. These data suggest therefore, that it would be advisable in future experiments to enrich for *VSG221*



non-expressers prior to plating cells. This has previously been used successfully in VSG switching experiments done by magnetic activated cell sorting (MACS) (Boothroyd *et al.*, 2009; Kim & Cross, 2010; Kim & Cross, 2011).

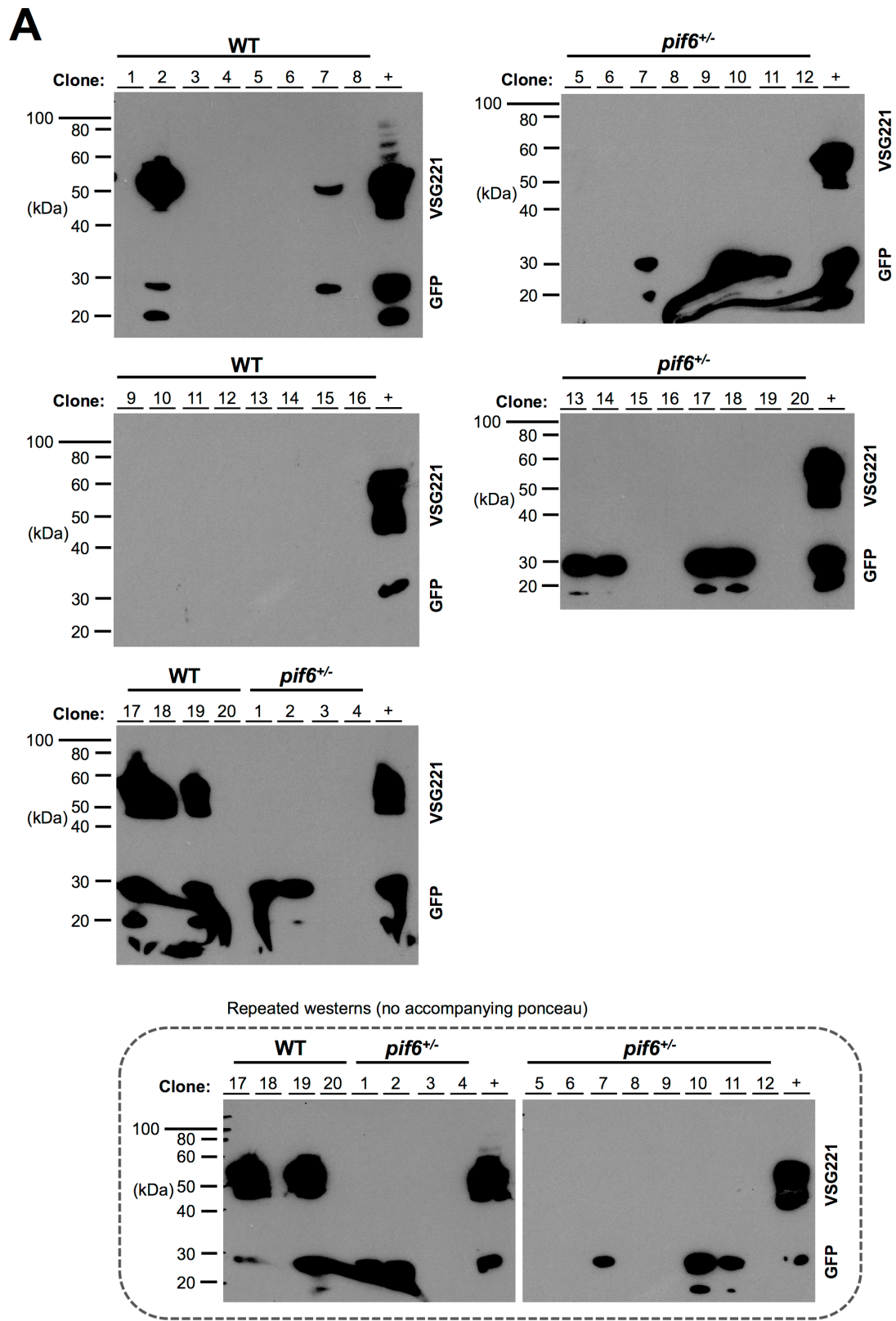


**Figure 4-38 Switching rate and switching profile of *pif6* mutants**  
**(A)** The switching rate of *GFP221hygTK* wild type and *GFP221hygTK pif6* mutants was calculated from the mean number of surviving clones, following treatment with ganciclovir after culture with or without puromycin. **(B)** The switching profiles of survivors of the different cell lines in the non-puromycin condition, represented as a percentage of total surviving clones analysed, determined using the PCRs and western blots described in the text. +P, with puromycin (1  $\mu\text{g.mL}^{-1}$ ); -P, without puromycin.

The high frequency of *TK* mutant cells in the non-puromycin condition is also of interest with respect to the VSG switching rate. The VSG switching rate of *pif6*<sup>-/-</sup> cells was elevated compared to wild type and remained elevated in the presence of puromycin (Fig. 4-38A). Whilst in *recq2*<sup>-/-</sup> mutants this was explained by dramatic shift to recombination and crossover events near the active VSG, the likely explanation here is the high percentage of *TK* mutants. The VSG switching rate of *pif6*<sup>-/-</sup> cells is roughly two-fold higher than wild type and 50% of *pif6*<sup>-/-</sup> survivors were classified as *TK* mutants, proportional to the increase in switching rate. These data suggest that the high percentage of *TK*

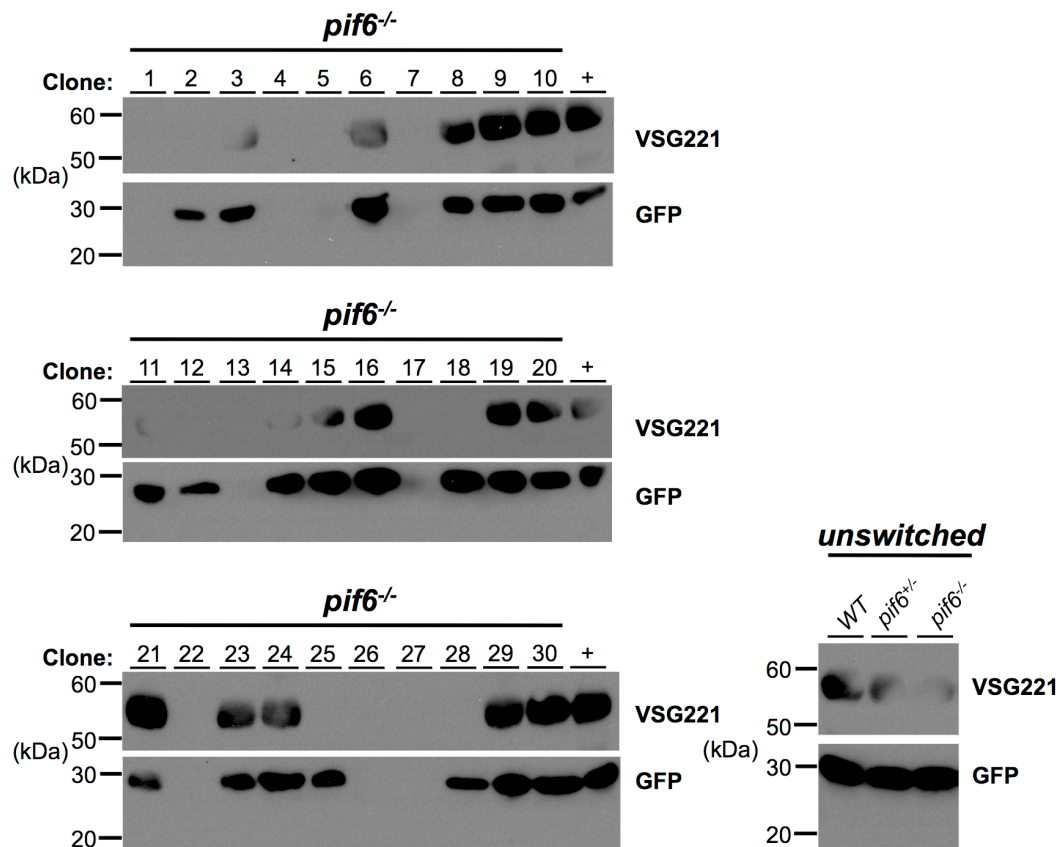


positive control for the presence of genomic DNA. Primers #156 and #157 were used to amplify a 522 bp region of VSG221. GFP221hygTK wild type genomic DNA (+) was used as a positive control in all PCRs. Distilled water was used as a negative control (-). -P, clones recovered from the assay conducted without puromycin selection; +P, clones recovered from the assay conducted with puromycin selection. PCR primer sequences were those as used in (Povelones *et al.*, 2012) and can be found in Appendix 7.1. Size markers (bp, ladder) are shown. PCRs were performed by Craig Lapsley.

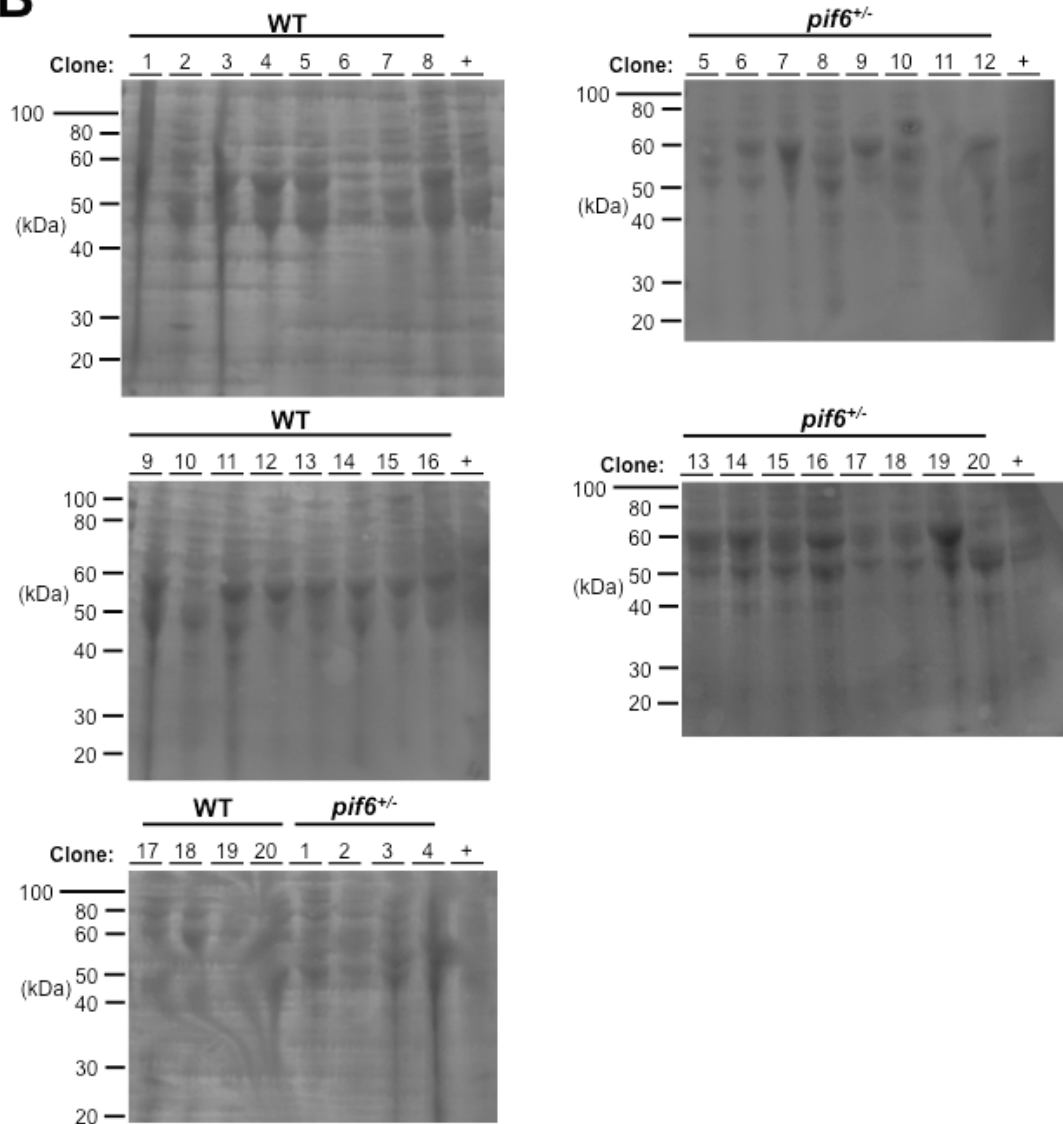


**Figure 4-40 Western analysis of *PIF6* VSG switching experiment**

(A) Western analysis of a selection of surviving GFP221hygTK wild type and *GFP221hygTK pif6* mutant clones. Total protein extracts were separated by SDS-PAGE and western blotted. Membranes were probed with rabbit anti-VSG221 antiserum (1:20,000 dilution, (Glover *et al.*, 2013a)) and rabbit anti-GFP antiserum (1:5000 dilution, abcam). Both membranes were probed with goat HRP-conjugated anti-rabbit antiserum (Life Technologies). -P, clones recovered from the assay conducted without puromycin selection; +P, clones recovered from the assay conducted with puromycin selection. GFP221hygTK wild type cells were used as a positive control (+). (B) Ponceau staining of the membranes shown in (A) after transfer. Membranes were incubated in ponceau solution for 10 minutes, washed and imaged. Size markers (kDa) are shown. Gaps and dashed lines indicate that lanes have been aligned in this figure after excision from multiple gels/membranes or from disparate parts of the same gel/membrane. Western blots of wild type and *pif6*<sup>+/+</sup> clones were performed by Craig Lapsley.

**Figure 4-40 continued**

**B**



**Figure 4-40 continued**

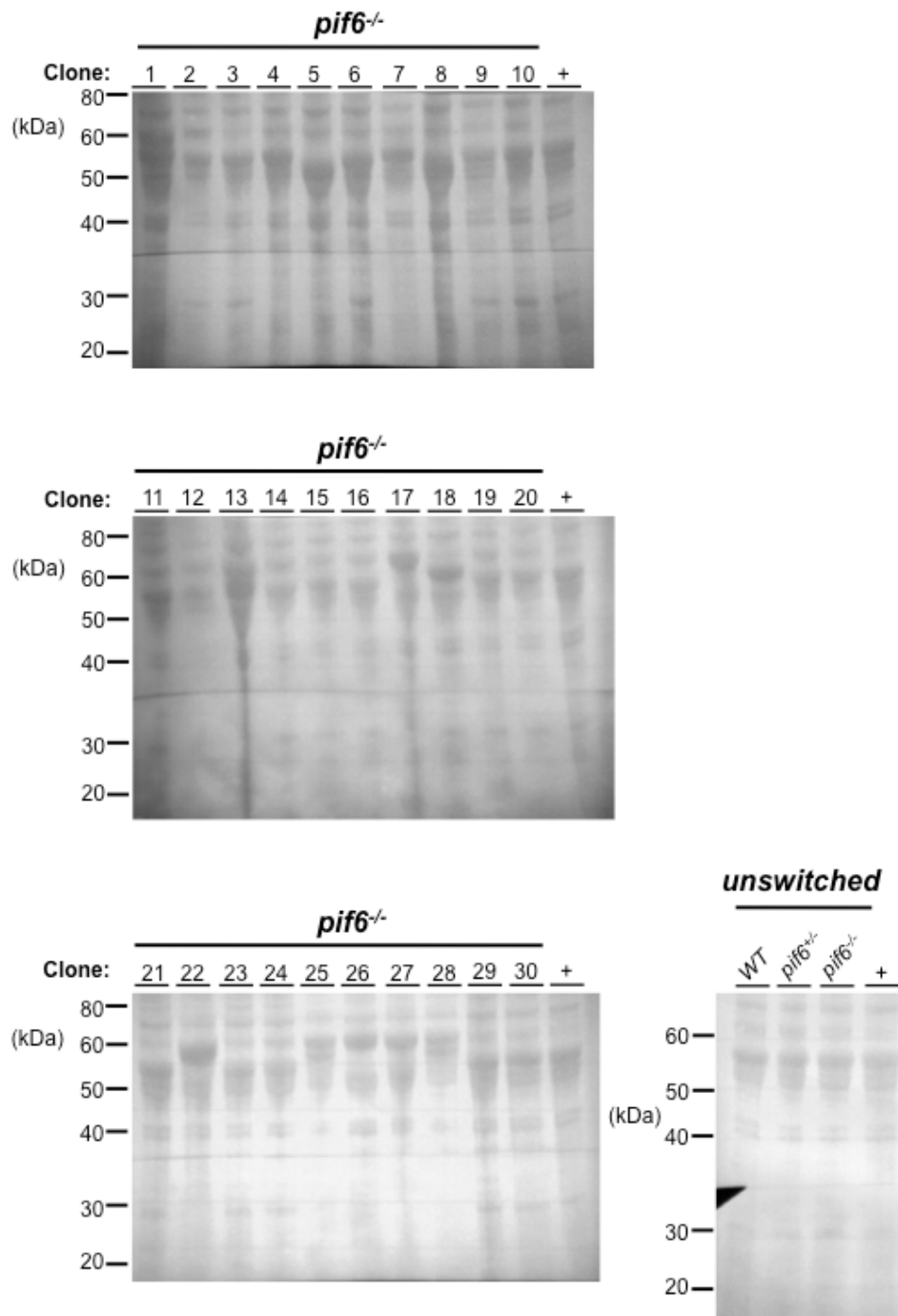


Figure 4-40 continued

| Cell line                  | Puromycin | Clone | PCR  |        | Western |     | Switching type            |
|----------------------------|-----------|-------|------|--------|---------|-----|---------------------------|
|                            |           |       | POLI | VSG221 | VSG221  | GFP |                           |
| WT                         | -         | 1     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 2     |      |        |         |     | TK mutant                 |
|                            |           | 3     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 4     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 5     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 6     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 7     |      |        |         |     | TK mutant**               |
|                            |           | 8     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 9     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 10    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 11    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 12    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 13    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 14    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 15    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 16    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 17    |      |        |         |     | TK mutant                 |
|                            |           | 18    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 19    |      |        |         |     | TK mutant                 |
|                            |           | 20    |      |        |         |     | ES GC or in situ + ES del |
| <i>pif6</i> <sup>+/-</sup> | -         | 1     |      |        |         |     | VSG GC                    |
|                            |           | 2     |      |        |         |     | Telomere XO               |
|                            |           | 3     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 4     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 5     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 6     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 7     |      |        |         |     | Telomere XO               |
|                            |           | 8     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 9     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 10    |      |        |         |     | Telomere XO               |
|                            |           | 11    |      |        |         |     | Telomere XO               |
|                            |           | 12    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 13    |      |        |         |     | Telomere XO               |
|                            |           | 14    |      |        |         |     | Telomere XO               |
|                            |           | 15    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 16    |      |        |         |     | in situ                   |
|                            |           | 17    |      |        |         |     | VSG GC                    |
|                            |           | 18    |      |        |         |     | Telomere XO               |
|                            |           | 19    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 20    |      |        |         |     | ES GC or in situ + ES del |
| <i>pif6</i> <sup>-/-</sup> | -         | 1     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 2     |      |        |         |     | Telomere XO               |
|                            |           | 3     |      |        |         |     | TK mutant                 |
|                            |           | 4     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 5     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 6     |      |        |         |     | TK mutant                 |
|                            |           | 7     |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 8     |      |        |         |     | TK mutant                 |
|                            |           | 9     |      |        |         |     | TK mutant                 |
|                            |           | 10    |      |        |         |     | TK mutant                 |
|                            |           | 11    |      |        |         |     | TK mutant                 |
|                            |           | 12    |      |        |         |     | Telomere XO               |
|                            |           | 13    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 14    |      |        |         |     | Telomere XO               |
|                            |           | 15    |      |        |         |     | TK mutant                 |
|                            |           | 16    |      |        |         |     | TK mutant                 |
|                            |           | 17    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 18    |      |        |         |     | VSG GC                    |
|                            |           | 19    |      |        |         |     | TK mutant                 |
|                            |           | 20    |      |        |         |     | TK mutant                 |
|                            |           | 21    |      |        |         |     | TK mutant                 |
|                            |           | 22    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 23    |      |        |         |     | TK mutant                 |
|                            |           | 24    |      |        |         |     | TK mutant                 |
|                            |           | 25    |      |        |         |     | VSG GC                    |
|                            |           | 26    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 27    |      |        |         |     | ES GC or in situ + ES del |
|                            |           | 28    |      |        |         |     | VSG GC                    |
|                            |           | 29    |      |        |         |     | TK mutant                 |
|                            |           | 30    |      |        |         |     | TK mutant                 |

 Negative


 Positive

Table 4-6 Summary of the analysis of clones from *PIF6* VSG switching experiment.

Summary of the data shown in Figs. 4-40 and 4-41. Telomere XO, telomere crossover; VSG GC, VSG gene conversion. Blue indicates a negative result, yellow indicates a positive result.

## 4.8 Summary

### 4.8.1 RECQ2 is involved in DSB repair and suppresses VSG switching

The clonal survival data presented here for *recq2*<sup>-/-</sup> mutants following I-SceI induction and DSB formation at the 70 bp repeats (HRES) and at a chromosomal internal location (HR1) showed that loss of this putative helicase resulted in ~50% lower survival than wild type, suggesting that RECQ2 is involved in DSB repair at both of these genomic locations. Repeats of these experiments did not replicate these data but analyses suggest these later data are less robust and that DSBs were not efficiently induced in these experiments, possibly due to loss of the I-SceI target sequence during culture. Further analysis of these cell lines to determine the presence or absence of the I-SceI recognition site would be necessary to confirm this. Additionally, these later repeats compromised analysis of the effect of deletion of a single *RECQ2* allele (*recq2*<sup>+/-</sup> cells).

A role for RECQ2 in DSB repair is consistent with proposed DNA repair roles seen in the context of increased sensitivity of *T. brucei recq2* mutants to phleomycin and MMS damage (see Section 3.4.3), both of which cause DSBs, the observed localisation of RECQ2 12myc to subnuclear foci after phleomycin treatment (see Section 3.9.1), and the role of other eukaryotic RecQ proteins in DSB repair. Efficient resection of DSB DNA ends has been demonstrated to require RECQ, in complex with DNA2 (Zhu *et al.*, 2008). For example, the mammalian RecQ protein BLM and *S. cerevisiae* SGS1 are required for normal DSB resection and mutants have increased sensitivity to DNA damaging agents and impaired HR-mediated repair (Gravel *et al.*, 2008; Langland *et al.*, 2002; Zhu *et al.*, 2008). RecQ proteins have been demonstrated to act at more than one point in HR; BLM plays an early role in DSB repair through DNA end resection and a later role in resolving recombination intermediates arising from HR-mediated repair (Mimitou & Symington, 2008). Sgs1 unwinds DNA at yeast DSBs, acting with the nuclease Dna2 and the ssDNA binding protein RPA, to resect DNA (Cejka *et al.*, 2010). Additionally, human RECQL4 mutants are less efficient in DSB repair in fibroblasts (Singh *et al.*, 2010).



Analysis of the VSG switching phenotype of *recq2* mutants demonstrated a two to three-fold increase in the VSG switching rate of *recq2*<sup>-/-</sup> mutants relative to wild type and *recq2*<sup>+/-</sup> cells, which was accounted for by a striking shift in the VSG switching profile of *recq2*<sup>-/-</sup> cells from ~60% VSG BES loss/GC events to ~90% VSG GC and telomere crossover switches. It seems likely that RECQ2 acts to suppress VSG switching by limiting crossover events, some of which may be associated with GC, in the active VSG BES, as has been proposed for TOPO3α and RMI1. Though *topo3α* and *rmi1* mutants have been reported to exhibit a greater elevation in VSG switching rate (10-40-fold and four-fold, respectively) than we see here for *recq2*<sup>-/-</sup>, the shift to a switching profile dominated by VSG GC and telomere crossover events (Kim & Cross, 2010; Kim & Cross, 2011) is strikingly similar.

Other RecQ proteins have been demonstrated to be important in suppressing hyper-recombination, including BLM and WRN (Prince *et al.*, 2001; Yamagata *et al.*, 1998). Suppression of mitotic hyper-recombination and resolution of recombination intermediates (double HJs, dHJ) to produce non-crossover products are functions also performed by the RTR complex in multiple eukaryotic organisms (Hartung *et al.*, 2008; Ira *et al.*, 2003; Raynard *et al.*, 2006; Wu & Hickson, 2003). Hyper-recombination is a phenotype of mutants of all three components of the RTR complex; human *blm* cells, *S. cerevisiae sgs1*, *top3* and *rmi1* mutants and *S. pombe rqh1* mutants all display hyper-recombination (Mullen *et al.*, 2005; Myung *et al.*, 2001; Stewart *et al.*, 1997; Traverso *et al.*, 2003; Ui *et al.*, 2001; Wallis *et al.*, 1989). Additionally, and as already discussed, *T. brucei rmi1* and *topo3* mutants also have elevated levels of recombination and crossover (Kim & Cross, 2010; Kim & Cross, 2011). The multiple functions RecQ proteins play in HR, dHJ resolution, Rad51 filament inhibition and D-loop formation, as well as DSB processing, are distinct (Bernstein *et al.*, 2010).

Taken together, the available data suggest that RECQ2 may function with RMI1 and TOPO3α as the RTR complex to suppress events in the vicinity of the VSG in the active expression site, limiting crossover events as well as gene conversions, and hence suppressing overall VSG switching rate. RMI1 and TOPO3α have been demonstrated to interact with one another (Kim & Cross, 2011) and co-immunoprecipitation of RECQ2 and RMI1/Topo3α and analysis of the VSG

switching phenotype of double mutants would enable this hypothesis to be tested.

The phenotypes observed above - impairment of *recq2* mutants in DSB repair at the 70 bp repeats of the active VSG BES, and their elevated VSG switching rate and altered profile - provide further insight into VSG switching. The data collected in this work are not consistent with the hypothesis that DSBs at the 70 bp repeats directly initiate VSG switching. As *recq2*<sup>-/-</sup> mutants are deficient in repair of DSBs at the 70 bp repeats of the active VSG BES (as well as in the interior of chromosome 11), it would be expected that the mutants would have a lower rate of VSG switching due to an inability to repair DSBs that form in this location (and whose repair would lead to VSG switching). Instead, the switching rate of *recq2*<sup>-/-</sup> mutants was two to three-fold higher than wild type, not lower. These genetic data suggest that the initiation of VSG switching cannot be accounted for by the direct formation of a DSB at the 70 bp repeats. Thus, although VSG switching is increasingly being modeled by I-SceI-directed DSB formation (Boothroyd *et al.*, 2009; Glover *et al.*, 2013a; Glover & Horn, 2014), this experimental model may be flawed. Though it is clear that induction of a DSB can elicit VSG switching, the analysis of RECQ2 function suggests this does not accurately mimic the natural route of switch initiation. Direct DSB formation in the active VSG BES may drive a distinct route of recombination than that normally used in VSG switching. Alternatively, it is possible that the formation of a DSB in the active VSG BES merely imposes selection for switchers through killing of cells that have not removed the I-SceI target sequence, since it is clear that the survival rate after DSB formation in this location is very low (such as compared with HR1). If so, this may suggest that DSBs are poorly recognised by the HR machinery in the VSG BES, perhaps explaining why no detectable RAD51 foci form in HRES cells (David Horn, personal communication) after I-SceI induction, but do so in HR1 (Glover *et al.*, 2008). Alternatively, homology searching, e.g. by RAD51, may be inefficient for breaks in the VSG BES.

Breaks, which are believed to be DSBs, have been detected in the active VSG BES (Boothroyd *et al.*, 2009; Glover *et al.*, 2013a). How these breaks arise and whether they are associated with VSG switching is not fully understood. One possibility for the source of these breaks is direct DSB formation via a site-specific endonuclease. For example, *S. cerevisiae* mating type switching uses

the HO endonuclease for formation of DSBs to initiate switching, reviewed in (Haber, 1998). However, cleavage by an endonuclease seems incompatible with the I-SceI data and furthermore, no such endonuclease has been identified and would have to be regulated such that it did not cleave the many other stretches of 70 bp repeats in locations other than active VSG BESs.

Replication stalling within the active VSG BES has also been proposed as the cause of breaks found in the active VSG BES (Kim & Cross, 2010; Kim & Cross, 2011). To date though, there is no evidence for replication in the active VSG BES, including the direction of replication or occurrence of replication stalling. RNAi knockdown of the replication factor ORC1/CDC6 (origin recognition complex 1/cell division cycle 6) causes an increase in VSG switching but this is primarily through VSG BES instability and *in situ* switches rather than recombination (Benmerzouga *et al.*, 2013). Inhibition of replication in *T. brucei* by treatment with HU causes an increase in the VSG switching rate (Shedder *et al.*, 2004). The possible involvement of replication stalling in the formation of breaks in the active VSG BES and VSG switching remains to be directly explored. To this end, single molecule analysis of replicated DNA (SMARD) would be a useful tool in attempting to determine of the direction of replication through the active VSG BES and whether replication regularly stalls in the region.

Several other hypotheses for the causes of DNA breaks in the active VSG BES have been proposed. Transcription-associated instability initiates immunoglobulin class switch recombination, whereby transcription directs the formation of breaks to initiate NHEJ, and ultimately affect class switching, reviewed by Stavnezer *et al.* (2008). The formation of unusual DNA structures has also been demonstrated to cause DNA breaks. The formation of a DNA/RNA hybrid at the *Neisseria gonorrhoeae* *PilE* (pilin expression) locus is necessary for the formation of a G4 DNA structure required for homologous recombination, leading to pilin antigenic variation (Cahoon & Seifert, 2013). Finally, pausing of the replication fork at the *S. pombe* *mat1* locus results in the formation of a DNA break and mating type switching by HR (Dalgaard & Klar, 2000; Kaykov *et al.*, 2004).

### 4.8.2 *mus81* mutants are deficient in DSB repair

The data presented here show that *mus81*<sup>-/-</sup> mutants are deficient in DSB repair both when such breaks are generated adjacent to the 70 bp repeats in the active VSG BES and at a chromosome internal location. As for RECQ2, this observation is consistent with the finding that *mus81*<sup>-/-</sup> mutants are hypersensitive to MMS, which causes replication stalling and DSBs, and phleomycin, which causes DNA SSBs and DSBs (see Section 3.4.3). Similar to *TbMUS81*, mitotic *mus81* mutants in *S. cerevisiae* and *S. pombe* are sensitive to MMS, though they are also sensitive to HU, to which *TbMUS81* does not appear sensitive (Bastin-Shanower *et al.*, 2003; Chang *et al.*, 2002; Doe *et al.*, 2004; Interthal & Heyer, 2000). Mus81 is not however required for DSB repair in vegetative yeast cells, as *mus81* mutants are not sensitive to bleomycin and ionizing radiation, both of which cause DSBs (Boddy *et al.*, 2000; Interthal & Heyer, 2000).

Mus81 acts as a heterodimer with Mms4 (methyl methanesulfonate sensitivity 4) in *S. cerevisiae* and Eme1 (essential meiotic structure-specific endonuclease 1) in *S. pombe* (Boddy *et al.*, 2001; Kaliraman *et al.*, 2001). The human partner protein for MUS81 is known as both Eme1 and Mms4 (Ciccia *et al.*, 2003; Ogrunc & Sancar, 2003). For simplicity, these MUS81-partner protein heterodimers are here referred to as MUS81\*, specifying the relevant organism(s).

MUS81\* plays a role in the completion of HR as a heterodimer with Eme1 (essential meiotic structure-specific endonuclease 1) in *S. pombe* and mammals, known as Mms4 in *S. cerevisiae*, to resolve recombination intermediates arising during replication (Boddy *et al.*, 2001; Interthal & Heyer, 2000; Kikuchi *et al.*, 2013) and is important in replication fork restart (Hanada *et al.*, 2007; Pepe & West). Mus81\* plays an essential role in meiosis in *S. pombe* and *S. cerevisiae*, resolving meiotic recombination intermediates to produce crossover products; *mus81* mutants display a lower rate of meiotic crossover (Boddy *et al.*, 2001; Kaliraman *et al.*, 2001; Osman *et al.*, 2003). There is evidence that MUS81\* acts on HJs and 3' flaps (Gaillard *et al.*, 2003) but MUS81/MMS4 shows a much greater activity with 3' flaps and replication structures than with HJs (Doe *et al.*, 2002; Kaliraman *et al.*, 2001; Pepe & West, 2014b) and the question of whether MUS81/MMS4 resolves intact HJs in *S. cerevisiae* has been a matter of debate, reviewed in Hollingsworth & Brill (2004).

The complicated picture of Mus81\* function in other eukaryotes makes it challenging to compare with the data on TbMus81 in this chapter. TbMus81 may be acting to resolve recombination intermediates arising from processes such as repair of DNA damage and replication fork restart. Failure to resolve these recombination intermediates in *mus81*<sup>-/-</sup> mutants may then cause the hypersensitivity to DSBs and DNA damaging agents. *mus81* mutants are synthetically lethal with an *sgs1* mutant in yeast, (Boddy *et al.*, 2000; Mullen *et al.*, 2001), suggesting that Mus81\* acts upon recombination intermediates generated by Sgs1 (Bastin-Shanower *et al.*, 2003). Analysis of whether a *recq2*<sup>-/-</sup>/*mus81*<sup>-/-</sup> double mutant is also lethal in *T. brucei* would be useful in determining whether TbMus81 performs a similar function in *T. brucei*. Given the strong phenotypes seen for *mus81* mutants in general repair, in DSB directed repair and in the localisation of the myc tagged protein to foci after damage, it will be interesting to see if and how it might contribute to VSG switching. This experiment was not performed in this thesis but the cell lines to conduct this experiment were generated (Section 4.7.3).

#### 4.8.3 PIF6: a complicated factor in *T. brucei* nuclear DNA repair

Similar to the data for RECQ2, the I-SceI survival data for PIF6 mutants gave inconsistent results upon repetition, though it has been argued that later repeats of the assay were less robust as it was unclear whether DSBs had been efficiently induced. It is possible that as with RECQ2, there was progressive loss of the I-SceI recognition sequence during culture. If we accept the data from the initial, and arguably most robust experiment, this suggest that whilst *pif6*<sup>-/-</sup> mutants are hypersensitive to DSBs at a chromosomal-internal location (HR1 cell line), they exhibit enhanced survival after DSB formation at the 70 bp repeats of the active VSG BES (HRES cell line), a response that is unexpectedly associated with retention of VSG221. If we rely on the later repeats of the clonal survival assays of *pif6* mutants following DSB induction in HRES, then loss of PIF6 does not cause an increase in *pif6*<sup>-/-</sup> survival compared to wild type. Nonetheless, *pif6*<sup>-/-</sup> survival was at least equivalent to wild type, indicating no impairment in DSB repair in the active VSG BES in *pif6*<sup>-/-</sup> mutants, a markedly different response to the strongly impaired survival after DSB formation at a chromosome-internal location.

The dramatically different response of *pif6*<sup>-/-</sup> mutants to DSB induction in these locations could be a result of how absence of PIF6 influences repair.

*S. cerevisiae* PIF1 suppresses telomere addition to DSBs and in its absence, DSBs are repaired by telomere addition and natural telomeres lengthen (Li *et al.*, 2014; Makovets & Blackburn, 2009; Schulz & Zakian, 1994). This repair mechanism would be advantageous to breaks in the vicinity of the 70 bp repeats of the active VSG BES due to fact that breaks here are telomere-proximal. Thus, only the VSG downstream of the DSB is lost - an effect that can be offset by VSG switching. In contrast, repair of a DSB located in the interior of chromosome 11 by addition of telomere repeats would create haploidy for many genes, which may be problematic. It would be interesting to investigate whether such processes occur in *pif6*<sup>-/-</sup> mutants, such as by measuring telomere length. This possibility does not fully explain the data, however, since retention of VSG221 was seen in survivors following DSB formation in HRES, an outcome that is incompatible with addition of a telomere at the break site. The expression and genomic location of VSG221 in these survivors is not known, however, which would be useful information in future experiments in order to determine how the DSB was repaired. Continued expression of and location of VSG221 in the same active VSG BES in the DSB survivors could, for example, suggest that MMEJ had been utilised to repair the break, limiting sequence loss to the region around the I-SceI site (including *PUR*). In contrast, VSG221 location in another expression site, and absence of expression, would indicate telomere crossover had been utilised. The altered survival of *pif6*<sup>-/-</sup> mutants after DSB formation in the nuclear genome is consistent with the phleomycin-induced localisation of PIF6 12myc, suggesting that the helicase acts in the nucleus. The putative increased survival of *pif6*<sup>-/-</sup> mutants in response to HRES breaks is also compatible with the observed improved survival of the same mutants in the presence of MMS. But what feature of the damage inflicted by MMS compared with a DSB in the VSG BES, and why no such altered repair phenotype is seen with phleomycin or HU damage, is unclear.

Following DSB induction, HR1 and HRES *pif6*<sup>-/-</sup> mutants undergo a G<sub>2</sub>/M cell cycle arrest, similar to wild type cells and as observed previously (Glover *et al.*, 2013a; Glover *et al.*, 2008). However, data presented here suggest some HRES *pif6*<sup>-/-</sup> cells may be in cell cycle arrest prior to I-SceI induction, possibly

due to leaky I-SceI expression. Furthermore, nuclear  $\gamma$ H2A accumulation, a marker of DNA damage (Canman, 2003; Glover & Horn, 2012; Polo & Jackson, 2011), appears to disappear more rapidly in HRES *pif6*<sup>-/-</sup> mutants than HRES wild type cells, suggesting that *pif6*<sup>-/-</sup> may repair the DSB more efficiently or alternatively, ignore the DNA damage and prematurely re-enter the cell cycle. The latter suggestion is supported by a slight increase in aberrant cells observed in HRES *pif6*<sup>-/-</sup> mutants at 24 hours post DSB induction, compared to virtually none in wild type.  $\gamma$ H2A expression and cell cycle analysis suggest that DSBs in a chromosome internal location (HR1) are repaired more efficiently;  $\gamma$ H2A expression decreased and cell cycle profiles returned to normal in wild type cells and *pif6* mutants at 24 hours. Again, these differences in HR1 and HRES suggest different roles of PIF6 in different genomic environments, but reconciling all the data is problematic.

The discussion above presents a complex picture of the role of PIF6 in *T. brucei* DSB repair, though it is clear that the helicase does provide a nuclear repair-associated function. The complexity of phenotypes is further added to by examining the effect of loss of PIF6 on VSG switching. The data generated in analysis of the VSG switching phenotype of *pif6* mutants was challenging to interpret due to the high rate of *TK* mutants generated and unusual switching profile of wild type cells compared to previous experiments. The switching rate of *pif6*<sup>-/-</sup> cells was higher than wild type but data suggested that this might be accounted for by the high rate of *TK* mutants generated in this experiment. To understand the role, if any, PIF6 has in VSG switching, it will be necessary to repeat this experiment, enriching for VSG221 non-expressers in order to exclude *TK* mutants from downstream analysis. The possibility that absence of PIF6 results in a higher rate of *TK* mutant generation could also be explored by analysis of survivors from both the + and - puromycin conditions. In summary, further work will be needed to try and understand how this enigmatic protein acts in *T. brucei* and to understand how the various repair and VSG switching phenotypes describe in *pif6* mutants (above) can be explained.

## **Chapter 5**

### **Analysis of the role of *T. brucei* RECQ1 in DNA repair**



## 5 Analysis of the role of *T. brucei* RECQ1 in DNA repair

### 5.1 Introduction

As discussed in Chapter 3, *T. brucei* encodes two RecQ like helicases, RECQ2 (Tb427.08.6690) and TbRECQ1 (Tb427.06.3580). The functions of the former in DNA repair and VSG switching were discussed in Chapters 3 and 4, which showed that TbRECQ2 is involved in both processes. TbRECQ2 plays a role in survival following induction of DSBs and following phleomycin or MMS treatment, as well as suppressing VSG switching by reducing crossover and VSG gene conversion events. In this chapter, the role of TbRECQ1 in DNA repair in *T. brucei* is investigated.

Mutations in the RecQ proteins WRN, BLM and RECQ4 are all associated with different disease syndromes in humans (reviewed in Croteau *et al.* (2014)), indicating specialisation of function among RecQ proteins in multicellular organisms containing more than one RecQ helicase. As TbRECQ1 and TbRECQ2 have so far not been characterised, it was unknown before starting this work whether they have overlapping or distinct functions. We therefore aimed to ascertain whether, like TbRECQ2, TbRECQ1 has a DNA repair function.

### 5.2 Sequence analysis of TbRECQ1

As described in Chapter 3 (Section 3.2) *TbRECQ1* (Tb427.06.3580) was identified and selected for analysis through BLAST searches using the protein sequences of eukaryotic RecQ-like proteins, searching against the *T. brucei* Lister 427 proteome on tritrypdb.org (Version 8.1). To begin to test the putative function and orthology of TbRECQ1, the predicted protein sequence was used for BLAST analysis.

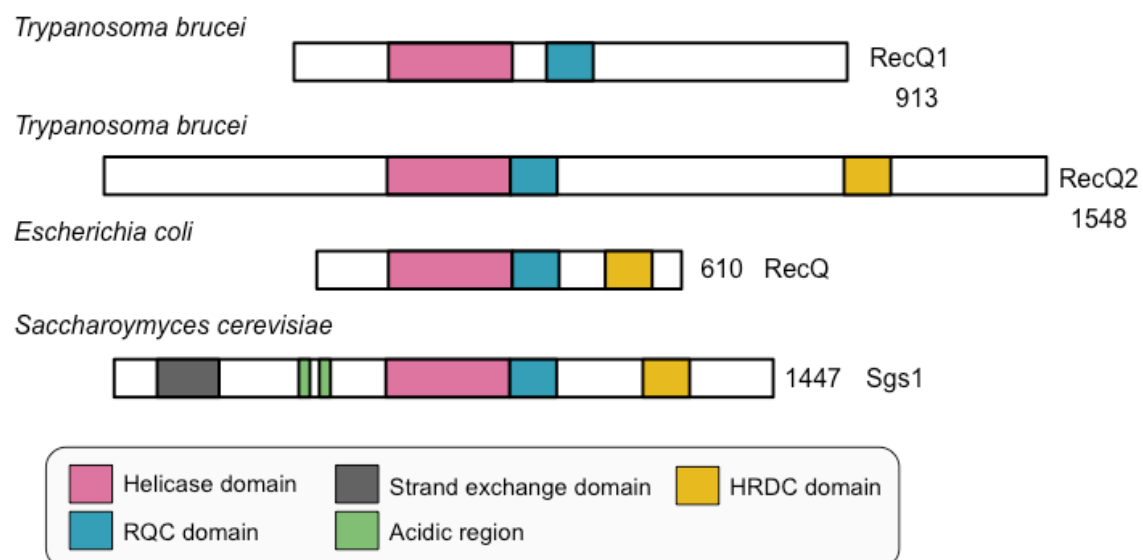
#### 5.2.1 Identification of TbRECQ1

BLAST analysis, using NCBI standard P-BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>; NCBI nr September 2014), was undertaken using the TbRECQ1 predicted amino acid sequence as the query sequence. The hits with the highest scores and lowest E values were all RecQ

helicases (Table 5-1). The top 16 hits were all *Trypanosoma spp* and *Leishmania spp* predicted RecQ-like proteins with expectancy values of 0. A number of bacterial RecQ-like proteins were retrieved as significant alignments with TbRECQ1, with E values in the range of  $10^{-167}$  -  $10^{-44}$ . In contrast to RECQ2, for which significant alignments with mammalian RecQ homologues were retrieved (Section 3.2.1), RECQ1 produced no strong alignments with mammalian RecQ proteins; significant alignments were almost exclusively bacterial (Table 5-1). BLAST searches were also performed (data not shown) using RecQ protein sequences from *H. sapiens*, *A. thaliana*, *C. elegans* and *S. cerevisiae* as the query sequences and using only *T. brucei* Lister 427 as the target organism (<http://tritryp.org>). TbRECQ1 (Tb427.06.3580) and TbRECQ2 (Tb427.08.6690) were consistently the only protein sequences retrieved with significant alignments. Additionally, several of the significant non-kinetoplastid TbRECQ1 BLAST results were used as query sequences in BLAST searches against *T. brucei* Lister 427 (<http://tritrypdb.org>). These non-kinetoplastid sequences were RecQ-like helicase sequences from *Haemophilus sputorum*, *Anditalea andensis* and *Vibrio brasiliensis* (protein sequences WP\_007523284.1, WP\_035071022.1 and WP\_006878004.1 respectively - <http://www.ncbi.nlm.nih.gov/protein>), and all returned TbRECQ1 and TbRECQ2 as the highest scoring sequences. In each of these searches, TbRECQ2 returned a stronger alignment than TbRECQ1.

#### Protein domain analysis searching InterproScan sequence search

(<http://www.ebi.ac.uk/interpro/search/sequence-search>) (Hunter *et al.*, 2012) was used to predict the domains present in TbRECQ1. TbRECQ1 is predicted to contain a DEAD/DEAH box helicase domain (approximately aa159-358) involved in nucleic acid and ATP binding, and a C-terminal helicase (RQC) domain (approximately aa407-602) found in helicases of multiple families (Linder, 2006) (Fig. 5-1). RQC domains are involved in protein-protein interactions (Bernstein *et al.*, 2010). Not predicted is an HRDC (helicase and RNaseD C-terminal) domain, found only in RecQ helicases, as well as RNaseD homologues. HRDC domains are thought to be involved in DNA binding (Bernstein *et al.*, 2010; Morozov *et al.*, 1997). Though HRDC domains are not present in all RecQ proteins (Bachrati & Hickson, 2003), the absence an HRDC domain in TbRECQ1 and its presence in TbRECQ2 suggests that TbRECQ1 may be more diverged from eukaryotic RecQ helicases than TbRECQ2.



**Figure 5-1 Representation of TbRECQ1 protein sequence**  
The predicted polypeptide sequence of TbRECQ1 (Tb427.06.3580) and comparison with that of TbRECQ2, *E. coli* RecQ and *S. cerevisiae* Sgs1. TbRECQ1 domains were predicted using InterproScan 5. Amino acid sequence length is shown on the right. Not to scale

| Description   | Protein Identifier<br>(ncbi.nlm.nih.gov/protein) | % Identity | Alignment<br>length<br>(residues) | E value   |
|---|--|------------|-----------------------------------|-----------|
| ATP-dependent DEAD/H DNA helicase recQ [Trypanosoma brucei TREU927]                                   | XP_845476.1                                      | 99.3       | 913                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ, putative [Trypanosoma brucei gambiense DAL972]                | XP_01177414.1                                    | 98.9       | 914                               | 0         |
| putative conserved ATP-dependent DEAD/H DNA helicase recQ famil [Trypanosoma vivax Y486]              | CCC48516.1                                       | 57.0       | 935                               | 0         |
| putative ATP-dependent DEAD/H DNA helicase recQ family [Trypanosoma grayi]                            | XP_009306147.1                                   | 55.8       | 948                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ family [Trypanosoma cruzi strain CL Brener]                    | XP_817577.1                                      | 55.1       | 940                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ family, putative [Trypanosoma cruzi]                           | EKG05656.1                                       | 55.2       | 940                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ family [Trypanosoma cruzi strain CL Brener]                    | XP_819079.1                                      | 55.2       | 940                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ family, putative [Trypanosoma cruzi marinkellei]               | EKF37818.1                                       | 54.7       | 931                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ family [Trypanosoma rangeli SC58]                              | ESL06759.1                                       | 54.8       | 934                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ family [Trypanosoma cruzi Dm28c]                               | ESS65314.1                                       | 60.4       | 594                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ family-like protein [Leishmania braziliensis MHOM/BR/75/M2904] | XP_001566851.1                                   | 41.5       | 915                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ family-like protein [Leishmania panamensis]                    | XP_010701248.1                                   | 41.4       | 910                               | 0         |
| TP-dependent DEAD/H DNA helicase recQ family-like protein [Leishmania mexicana MHOM/GT/2001/U1103]    | XP_003877373.1                                   | 40.9       | 900                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ family-like protein [Leishmania major strain Friedlin]         | XP_001684834.1                                   | 41.4       | 902                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ family-like protein [Leishmania infantum JPCM5]                | XP_001467076.1                                   | 41.6       | 899                               | 0         |
| ATP-dependent DEAD/H DNA helicase recQ family-like protein [Leishmania donovani]                      | XP_003862942.1                                   | 41.5       | 899                               | 0         |
| unnamed protein product [Phytomonas sp. isolate Hart1]  | CCW68096.1                                       | 43.4       | 746                               | 0         |
| ATP-dependent DNA helicase RecQ [Argomonas deanei]  | EPY21020.1                                       | 40.2       | 734                               | 6.78E-167 |
| ATP-dependent DEAD/H DNA helicase recQ family- like protein [Leptomonas seymouri]                     | KPI90195.1                                       | 43.3       | 607                               | 5.29E-142 |
| ATP-dependent DEAD/H DNA helicase recQ family- like protein [Leptomonas seymouri]                     | KPI90195.1                                       | 43.0       | 235                               | 2.84E-46  |
| unnamed protein product [Trypanosoma congolense IL3000]   | CCC91067.1                                       | 61.1       | 332                               | 1.65E-129 |
| unnamed protein product [Phytomonas sp. isolate EM1]  | CCW61605.1                                       | 44.2       | 480                               | 2.94E-111 |
| unnamed protein product [Phytomonas sp. isolate EM1]  | CCW61605.1                                       | 44.5       | 319                               | 4.16E-65  |
| PREDICTED: ATP-dependent DNA helicase Q4 [Trichogramma pretiosum]                                     | XP_014234792.1                                   | 33.7       | 412                               | 5.29E-54  |
| PREDICTED: ATP-dependent DNA helicase Q4 [Ceratosolen solmsi marchali]                                | XP_011502558.1                                   | 32.5       | 496                               | 2.07E-53  |
| ATP-dependent DNA helicase RecQ [Haemophilus sputorum]  | WP_007523284.1                                   | 32.4       | 435                               | 9.98E-51  |
| ATP-dependent DNA helicase RecQ [Haemophilus parahaemolyticus]  | WP_005705706.1                                   | 33.2       | 440                               | 1.58E-50  |
| putative ATP-dependent DEAD/H DNA helicase recQ family [Trypanosoma cruzi Dm28c]                      | ESS65315.1                                       | 40.5       | 289                               | 2.25E-50  |

**Table 5-1 Top protein TbRECQ1 BLAST hits**  
**Top BLAST hits using the TbRECQ1 predicted amino acid sequence**

|  |                 |      |     |          |
|--|-----------------|------|-----|----------|
| hypothetical protein [ <i>Bacillus aurantiacus</i> ]                                       | WP_026689447.1  | 31.9 | 427 | 1.76E-49 |
| ATP-dependent DNA helicase RecQ [ <i>Haemophilus paraprohaemolyticus</i> ]                 | WP_005708718.1  | 32.0 | 440 | 2.49E-49 |
| ATP-dependent DNA helicase RecQ [ <i>Anditalea andensis</i> ]                              | WP_035071022.1  | 29.4 | 425 | 4.33E-49 |
| ATP-dependent DNA helicase RecQ [ <i>Vibrio brasiliensis</i> ]                             | WP_006878004.1  | 31.6 | 443 | 5.57E-49 |
| ATP-dependent DNA helicase RecQ [ <i>Bifidobacterium thermacidophilum</i> ]                | WP_052400980.1  | 32.3 | 436 | 8.72E-49 |
| ATP-dependent DNA helicase RecQ [ <i>Actinobacillus equuli</i> ]                           | WP_039197383.1  | 32.4 | 432 | 1.27E-48 |
| ATP-dependent DNA helicase RecQ [ <i>Pseudomonas syringae</i> pv. aptata str. DSM 50252]   | EGH78679.1      | 33.2 | 413 | 1.84E-48 |
| PREDICTED: ATP-dependent DNA helicase Q4 isoform X3 [ <i>Chinchilla lanigera</i> ]         | XP_005395970.1  | 32.6 | 466 | 2.04E-48 |
| ATP-dependent DNA helicase RecQ [ <i>Bifidobacterium thermacidophilum</i> ]                | IWP_044092856.1 | 31.9 | 436 | 2.09E-48 |
| ATP-dependent DNA helicase RecQ [ <i>Cyclobacterium amurskyense</i> ]                      | WP_048642395.1  | 29.0 | 420 | 2.76E-48 |
| ATP-dependent DNA helicase RecQ [ <i>Bifidobacterium thermacidophilum</i> subsp. porcinum] | KFI99497.1      | 32.2 | 425 | 3.00E-48 |
| PREDICTED: ATP-dependent DNA helicase Q4 isoform X1 [ <i>Chinchilla lanigera</i> ]         | XP_005395966.1  | 32.6 | 466 | 3.77E-48 |
| ATP-dependent DNA helicase, RecQ family [ <i>Thermobaculum terrenum</i> ATCC BAA-798]      | ACZ42826.1      | 33.7 | 416 | 4.67E-48 |
| ATP-dependent DNA helicase RecQ [ <i>Cyclobacterium qasimii</i> ]                          | WP_020892691.1  | 28.8 | 420 | 4.92E-48 |
| PREDICTED: ATP-dependent DNA helicase Q4 isoform X4 [ <i>Nannospalax galii</i> ]           | XP_008832617.1  | 33.0 | 469 | 8.06E-48 |
| ATP-dependent DNA helicase domain protein [ <i>Pseudomonas caricapapayae</i> ]             | KPW55436.1      | 32.9 | 413 | 8.57E-48 |
| ATP-dependent DNA helicase RecQ [ <i>Rhodobacteraceae bacterium</i> HLUCCO18]              | KPQ17774.1      | 32.3 | 455 | 8.58E-48 |
| hypothetical protein [ <i>Thermobaculum terrenum</i> ]                                     | WP_049823046.1  | 33.7 | 416 | 1.43E-47 |
| PREDICTED: ATP-dependent DNA helicase Q4 isoform X1 [ <i>Nannospalax galii</i> ]           | XP_008832614.1  | 33.0 | 469 | 1.90E-47 |
| PREDICTED: ATP-dependent DNA helicase Q4 isoform X2 [ <i>Nannospalax galii</i> ]           | XP_008832615.1  | 33.0 | 469 | 1.94E-47 |
| ATP-dependent DNA helicase RecQ [ <i>Actinobacillus capsulatus</i> ]                       | WP_018652670.1  | 30.9 | 447 | 2.12E-47 |
| ATP-dependent DNA helicase Q4 [ <i>Orycteropus afer</i> ]                                  | XP_007954294.1  | 31.6 | 493 | 2.69E-47 |

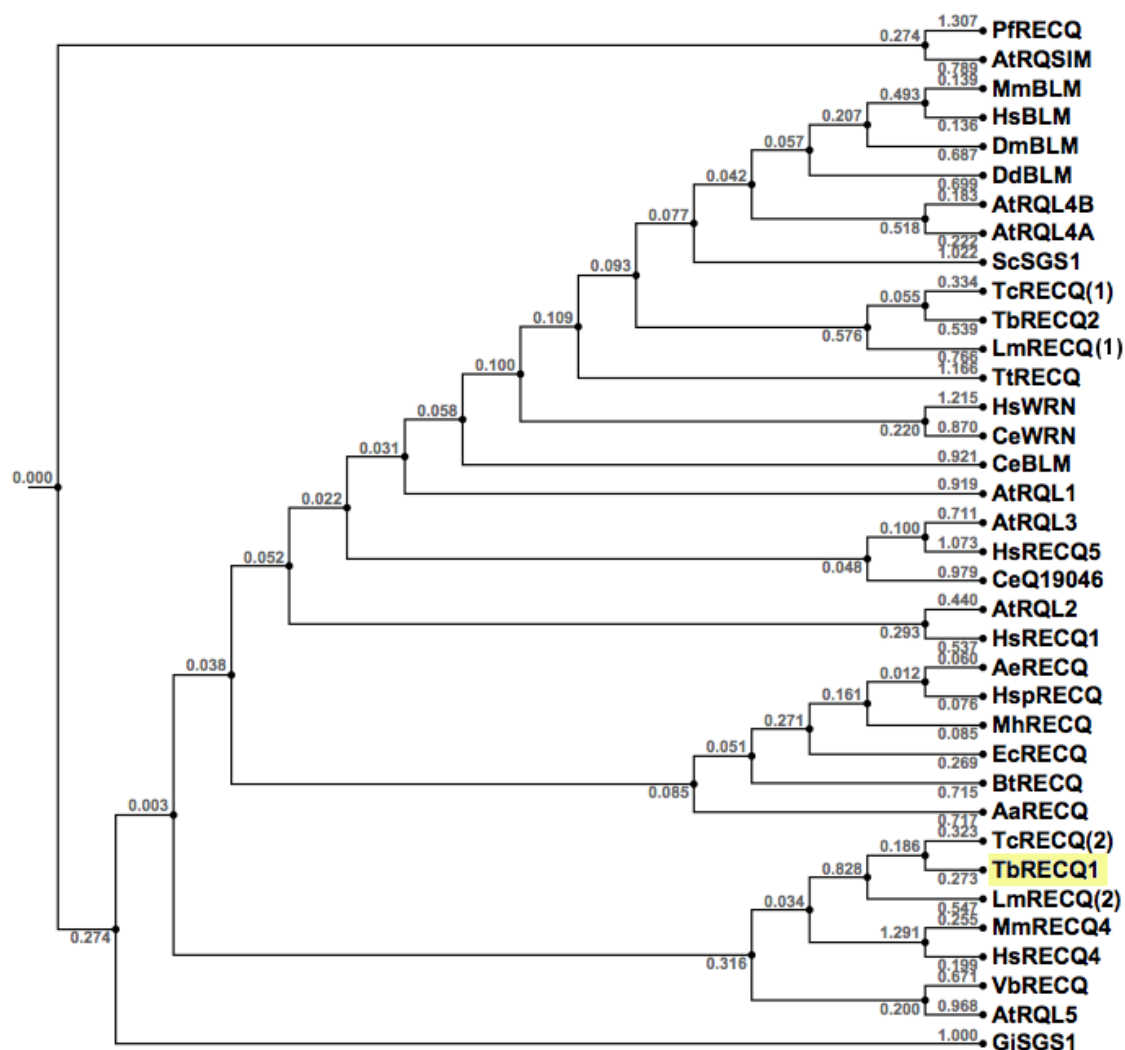
Table 5-1 continued

### 5.2.2 Phylogenetic analysis

A multiple sequence alignment of the complete predicted sequences of TbRECQ1 and TbRECQ2 with other eukaryotic RecQ proteins in Chapter 3 (Section 3.2.1) showed that TbRECQ2 is more similar to eukaryotic RecQ proteins than TbRECQ1. For this reason, and the finding above that most of the TbRECQ1 BLAST results were bacterial RecQ helicases, TbRECQ1 was aligned with bacterial RecQ helicases in addition to eukaryotic RecQ helicases and a phylogenetic tree was constructed from this alignment using CLC Genomics Workbench 6.

This tree (Fig. 5-2) shows that despite the dominance of bacterial RecQ-like proteins retrieved in BLAST analysis, TbRECQ1 aligns more closely to eukaryotic RecQ-like proteins. The retrieval of predominantly bacterial RecQ-like proteins in BLAST analysis is therefore most likely due to (1) TbRECQ1 being more diverged from eukaryotic RecQ-like proteins, hence aligning to RecQ helicases from a wider range of organisms and (2) the large number of bacterial genomes present in the NCBI protein database.

TbRECQ1 groups with other RecQ-like proteins, demonstrating together with the BLAST analysis that TbRECQ1 is a RecQ-like protein. However, TbRECQ1 clusters separately from most eukaryotic RecQ-like proteins, whereas TbRECQ2 groups more closely with eukaryotic RecQ-like proteins. This indicates that TbRECQ1 is more diverged from most eukaryotic RecQ-like helicases than TbRECQ2. TbRECQ1 also groups very closely with one of the two putative RecQ-like proteins from each of *T. cruzi* and *L. major*, while the second putative RecQ-like helicase from each of these three kinetoplastids groups with the eukaryotic helicases (yeast SGS1 and mammalian BLM homologues) (Fig. 5-2). This again suggests that TbRECQ1 may be more diverged from most eukaryotic RecQ-like proteins, perhaps due to evolving to perform a specialised, kinetoplastid-specific function. If so, such a specialised function might be shared by the *L. major* and *T. cruzi* putative RecQ-like proteins that group with TbRECQ1.



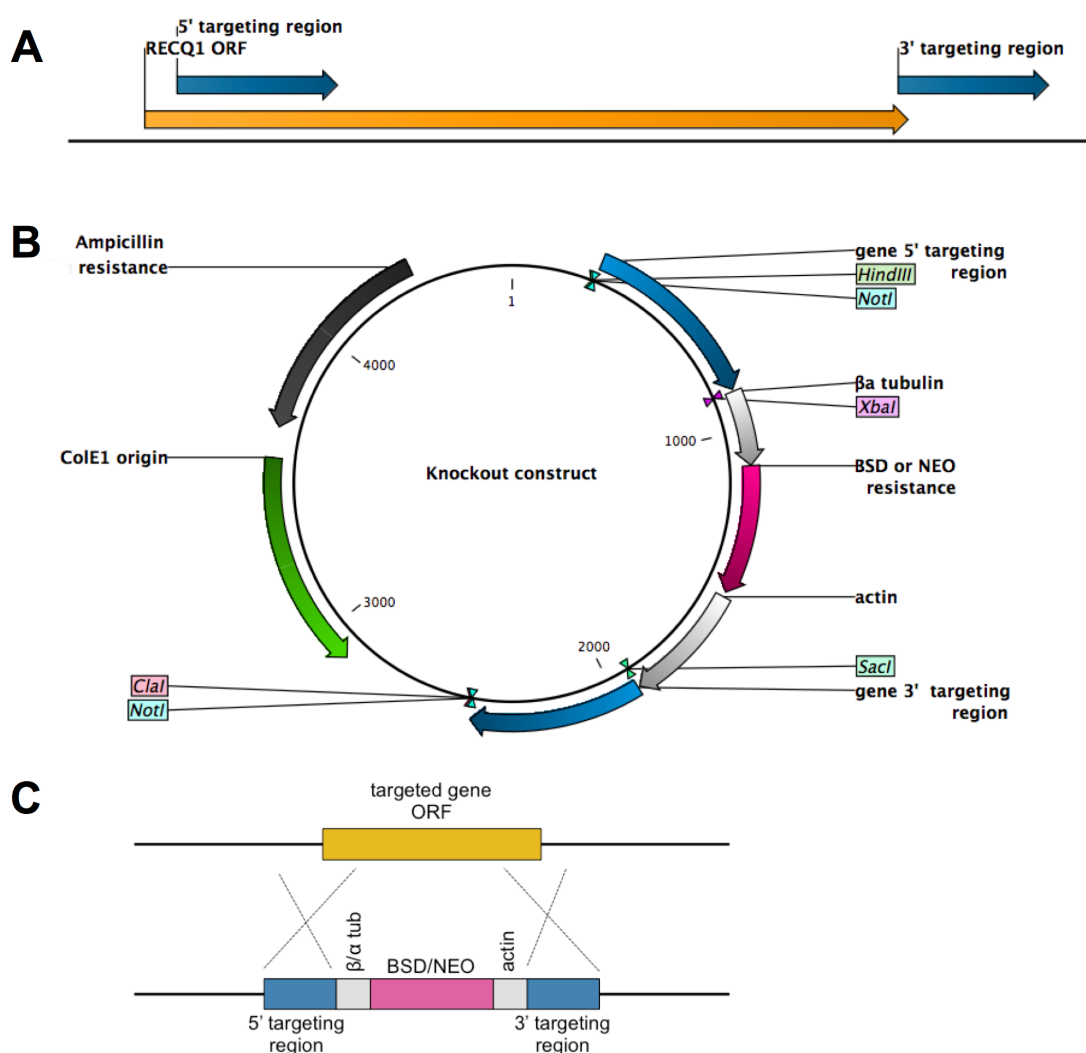
**Figure 5-2 Neighbour-joining tree of TbRECQ1 and other RecQ-like helicases**  
A tree was constructed from the alignment of TbRECQ1, eukaryotic and bacterial RecQ-like helicase protein sequences using CLC Genomics Workbench 6. At, *A. thaliana*; Dd, *Dictyostelium discoideum*; Mm, *Mus musculus*; Ce, *C. elegans*; Sc, *S. cerevisiae*; Dm, *D. melanogaster*; Tc, *T. cruzi*; Tb, *T. brucei*; Lm, *L. major*; Tt, *Tetrahymena thermophilus*; Hs, *H. sapiens*; Gi, *Giardia intestinalis*; Pf, *P. falciparum*. The two putative RecQ-like helicases in *T. cruzi* and *L. major* are indicated by (1) and (2) in the sequence name, assigned arbitrarily, while the *T. brucei* RecQ-like proteins follow the nomenclature used in this thesis (TbRECQ1 and TbRECQ2). See Appendix 7.2 for gene accession numbers for sequences used.

## 5.3 Generation of a RECQ1 RNAi cell line

### 5.3.1 Attempts at generating a RECQ1 knockout cell line

To investigate the function of TbRECQ1 and specifically whether it is involved in DNA repair, it was aimed to generate *RECQ1* heterozygous (+/-) and homozygous (-/-) mutants. The knockout constructs used in the attempt to generate *recq1*<sup>+/-</sup> and *recq1*<sup>-/-</sup> mutants used the same plasmid backbone and strategy as that used to generate *recq2*, *mus81* and *pif6* mutants (see Section 3.3). A 580 bp region of

the 5' end of the *RECQ1* ORF and 544 bp region of the 3' end of the *RECQ1* ORF and the 3' UTR, depicted in Figure 5-3A, were amplified by PCR and cloned into the knockout construct, flanking a blasticidin (*BSD*) or neomycin (*NEO*) resistance cassette. PCR-specific details are listed in Figure 5-3. A generalised schematic of the knockout plasmids used is shown in Figure 5-3B. Knockout constructs were linearised using *NotI* and integration into the genome and replacement of the targeted gene was facilitated by the 5' and 3' targeting regions of the linearised knockout plasmid (Fig. 5-3C).



**Figure 5-3** Schematic of *RECQ1* knockout constructs and knockout strategy (A) A 580 bp region (blue) of the 5' end of the *RECQ1* ORF (orange) and 544 bp region (blue) of the 3' end of the *RECQ1* ORF and the *RECQ1* 3' UTR was amplified by PCR. These were the targeting regions in the knockout constructs to knockout *RECQ1*. PCR primer sequences: *RECQ1* 5' targeting region primers #26 & #27, *RECQ1* 3' targeting region primers #28 & #29. Primer sequences can be found in Appendix 7.1. (B) Schematic of the knockout constructs. 3' and 5' targeting regions of *RECQ1* (blue) flank the resistance cassette containing a neomycin (*NEO*) or blasticidin (*BSD*) resistance gene (pink). *NotI* (as indicated) was used to linearise the constructs and restriction enzyme sites are indicated. Sizes shown (bp), length of the 5' and 3' targeting fragments were specific to the gene targeted. (C) Knockout strategy used to generate heterozygous (+/-) and knockout (-/-)



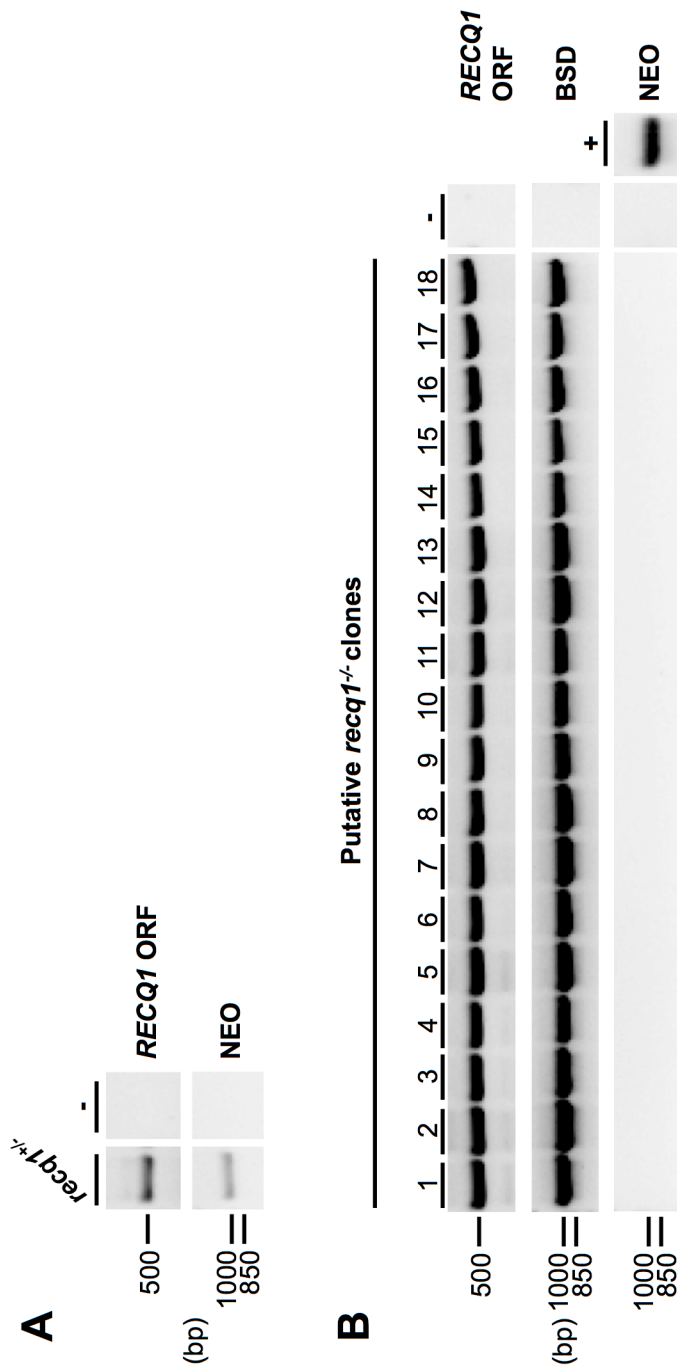
mutants. Generalised representation of the linearised knockout *BSD* and *NEO* resistance constructs (bottom) relative to the targeted gene (ORF, orange) and its flanking regions (black line). The 5' and 3' targeting regions in the constructs correspond to the 5'-most or 3'-most regions of the *RECQ1* ORF and its flanking regions.  $\beta/\alpha$  tub,  $\beta/\alpha$  tubulin; actin IR, actin intergenic region; *BSD*, blasticidin resistance; *NEO*, G418 resistance. Not to scale.

The same PCRs as used in Chapters 3 and 4 (see Sections 3.3.3 and 4.3.2) were used to test for successful integration of these constructs in transformed clones. Briefly, three PCRs were used. The first, “*RECQ1* ORF”, amplified a 226 bp region of the *RECQ1* ORF to test absence of the *RECQ1* ORF in putative *recq1*<sup>-/-</sup> clones (primers #75 and #76). The second PCR amplified a region using a forward primer lying upstream of the 5' targeting fragment present in the knockout constructs (primer #50) and a reverse primer lying in the *BSD* (primer #154) or *NEO* ORFs (primer #155), 1077 bp for *BSD* and 1085 bp for *NEO*. The “*BSD*” and “*NEO*” PCRs confirmed the integration of the respective knockout constructs in the correct location, due to the forward primer (#50) binding upstream of the 5' UTR region present in the construct.

Wild type Lister 427 cells transformed with the *RECQ1:: $\Delta$ BSD* construct successfully generated *recq1*<sup>+/-</sup> clones following selection with 10  $\mu\text{g}.\text{mL}^{-1}$  blasticidin, confirmed by PCR as described above (data not shown). However, repeated attempts at generation of *recq1*<sup>-/-</sup> cells by transformation of *recq1*<sup>+/-</sup> with the  *$\Delta$ RECQ1::NEO* construct and selection using 5  $\mu\text{g}.\text{mL}^{-1}$  blasticidin and 2.5  $\mu\text{g}.\text{mL}^{-1}$  G418 were unsuccessful. Surviving clones were shown not to be *recq1*<sup>-/-</sup> using the diagnostic PCRs described above (data not shown). In some cases, no double resistant clones were recovered, whereas in others cases double resistant clones retained the *RECQ1* ORF.

The above data, in contrast to the efficiency with which *recq2*<sup>-/-</sup> mutants were generated, suggested that *RECQ1* may be essential in BSF *T. brucei*. However, the transformations were repeated in reverse, transforming the  *$\Delta$ RECQ1::NEO* construct first, followed by the  *$\Delta$ RECQ1::BSD* construct, to ensure that the  *$\Delta$ RECQ1::BSD* construct could be integrated into wild type cells. Altered selective drug concentrations were also used (described below). Wild type Lister 427 cells were transformed with the  *$\Delta$ RECQ1::NEO* construct and antibiotic resistant clones were selected with 5  $\mu\text{g}.\text{mL}^{-1}$  G418. Fifteen putative *recq1*<sup>+/-</sup> clones were obtained, of which 13 were shown to contain the correctly integrated  *$\Delta$ RECQ1::NEO* construct (data not shown). One of these clones (Fig.

5-4A) was taken forward for transformation of the  $\Delta RECQ1::BSD$  construct. Given the failure to obtain antibiotic *recq1*<sup>-/-</sup> clones in previous attempts, antibiotic resistant cells were selected using two different conditions: 5  $\mu\text{g.mL}^{-1}$  blasticidin and 2.5  $\mu\text{g.mL}^{-1}$  G418, or 10  $\mu\text{g.mL}^{-1}$  blasticidin alone. No surviving clones were obtained following selection with G418 and blasticidin. Eighteen putative *recq1*<sup>-/-</sup> clones were obtained following selection with blasticidin only, all of which still contained the *RECQ1* ORF (Fig. 5-4B) indicating that none of these clones were *recq1*<sup>-/-</sup> mutants. In all of these clones, the *BSD* knockout construct had integrated correctly but the *NEO* knockout construct had been lost (Fig. 5-4B). As deletion of the *RECQ1* ORF was unsuccessful, this provided further evidence that *RECQ1* is essential in BSF *T. brucei*.



**Figure 5-4** PCR analysis of putative *recq1<sup>-/-</sup>* clones. Agarose gels of PCR products generated from genomic DNA of (A) *recq1<sup>-/-</sup>* cells and (B) putative *recq1<sup>-/-</sup>* clones, using primers described in the text. **A** 226 bp region of the *RECQ1* ORF (primers #75 & #76) was PCR amplified ("RECQ1 ORF"). Integration of the blasticidin (*BSD*) knockout construct was tested by PCR amplification using primers #50 & #154 (1077 bp PCR product) that bind in the *BSD* resistance cassette and upstream of the *RECQ1* 5' targeting region used in the knockout construct ("BSD"). Integration of the neomycin (*NEO*) knockout construct was tested by PCR amplification using primers #50 and #155 (1085 bp PCR product) that bind in the *NEO* resistance gene and downstream of the *RECQ1* 3' targeting region used in the knockout construct ("NEO"). Primer sequences can be found in Appendix 7.1. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels or from disparate parts of the same gel; size markers are shown (ladder, bp).

### 5.3.2 Generation of a RECQ1 RNAi construct

The inability to successfully obtain *recq1*<sup>-/-</sup> cells suggested that *RECQ1* is essential in BSF form *T. brucei*. To test this and to investigate the role of *RECQ1*, it was next aimed to generate an inducible *RECQ1* RNAi line.

#### 5.3.2.1 RNAi in *T. brucei*

RNAi, first reported by Fire *et al.* (1998), is a process whereby double stranded RNA (dsRNA) molecules are cleaved to generate short interference RNA (siRNA) molecules, which after processing, bind to their target mRNAs, leading to destruction of the target mRNA and subsequent inhibition of synthesis of the corresponding protein.

*T. brucei* was the first protozoan parasite demonstrated to possess a functional RNAi system (Ngo *et al.*, 1998). RNAi has since been successfully developed and widely utilised as an experimental tool to study gene function in *T. brucei*, reviewed in Kolev *et al.* (2011). Sense and anti-sense RNA transcripts are processed by two Dicer-like enzymes, TbDCL1 and TbDCL2, located in the cytoplasm and nucleus respectively, into dsRNA siRNA molecules 20-30 nucleotides in length (Patrick *et al.*, 2009; Shi *et al.*, 2006). siRNAs are loaded into the TbAGO1, the single *T. brucei* Argonaute protein (Durand-Dubief & Bastin, 2003; Shi *et al.*, 2004), to form the RNAi induced silencing complex (RISC). TbAGO1, together with TbRIF4 and TbRIF5 (RNA interference factor) processes siRNA duplexes into the mature, single stranded form (Barnes *et al.*, 2012). One strand of the duplex is removed (known as the passenger strand) and the remaining strand (known as the guide strand) interacts directly with Argonaute to target the mRNA, base pairing with it, which leads to mRNA cleavage and consequent down regulation of the corresponding protein.

#### 5.3.2.2 Gateway RNAi system and cloning of *RECQ1* RNAi construct

A number of genetic tools have been developed to take advantage of the *T. brucei* RNAi system. Inducible expression of dsRNA molecules homologous to a target gene results in degradation of mRNA and down regulation of synthesis of the target gene's corresponding protein. The system employed here is the same as that developed by Jones *et al.* (2014). This system uses the 2T1 *T. brucei* BSF

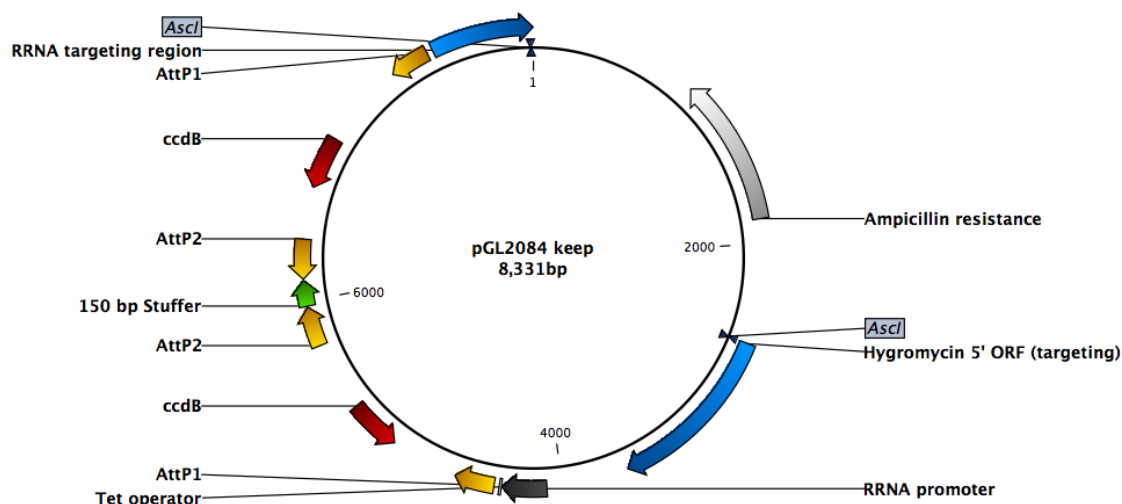
cell line, developed by Alsford *et al.* (Alsford *et al.*, 2005b). The 2T1 cell line (Fig. 5-6B) expresses TetR (tetracycline repressor) from the pHD1313 cassette integrated at the tubulin locus; see (Alibu *et al.*, 2005). Additionally, the 2T1 cell line contains a 3' hygromycin (*HYG*) resistance gene fragment driven by an *EP1* procyclin promoter upstream of a puromycin resistance (*PUR*) ORF driven by a regulated ribosomal RNA (*RRNA*) promoter, integrated at a ribosomal spacer locus on chromosome 2a. The 3' *HYG* and *RRNA* sequences, which flank the *PUR* gene, create sequence homology to facilitate integration of linearised constructs flanked by the same sequences.

2T1 cells are puromycin resistant and hygromycin sensitive. Transformation of 2T1 cells with a pRPa construct (Alsford *et al.*, 2005b), which contains RNAi targeting sequences flanked by a 5' *HYG* ORF and *RRNA* sequence, results in stable integration of the construct at the tagged *RRNA* spacer locus, loss of *PUR* and reconstitution of the *HYG* ORF. Transformed cells should thus be puromycin sensitive and hygromycin and phleomycin resistant (phleomycin resistance being conferred by the TetR expression construct).

The RNAi system used by Jones *et al.* (2014) utilises a modified version of the pRPa<sup>ISL</sup> construct (Alsford & Horn, 2008), pGL2084 (Fig. 5-5), to transform 2T1 cells. pGL2084 contains a Gateway recombination cassette located between the *HYG* 5'-*RRNA* promoter and *RRNA* targeting sequences. This cassette consists of two pairs of inverted AttP sites from the Gateway cloning system (Life Technologies). In a Gateway reaction, AttB sites recombine with AttP sites, catalysed by a "BP clonase" enzyme, such that a single PCR product flanked by AttB sites is recombined into the pGL2084 plasmid at the site of the AttP sequences, forming an inverted tandem repeat.

A 150 bp stuffer fragment is located between the pairs of AttP repeats, which allows the formation of a hairpin between the inverted PCR sequences once the construct is transcribed into RNA. pGL2084 also contains the *ccdB* gene between both sets of AttP repeats. The product of the *ccdB* gene is lethal in bacteria. However, upon successful recombination of the AttB-flanked PCR product into pGL2084, both *ccdB* sequences are lost and the plasmid can be propagated. pGL2084 also contains the *TetO* (Tet operator) sequence, as in pRP<sup>ISL</sup>, which

results in the repression of transcription of the sequences downstream in the absence of tetracycline.

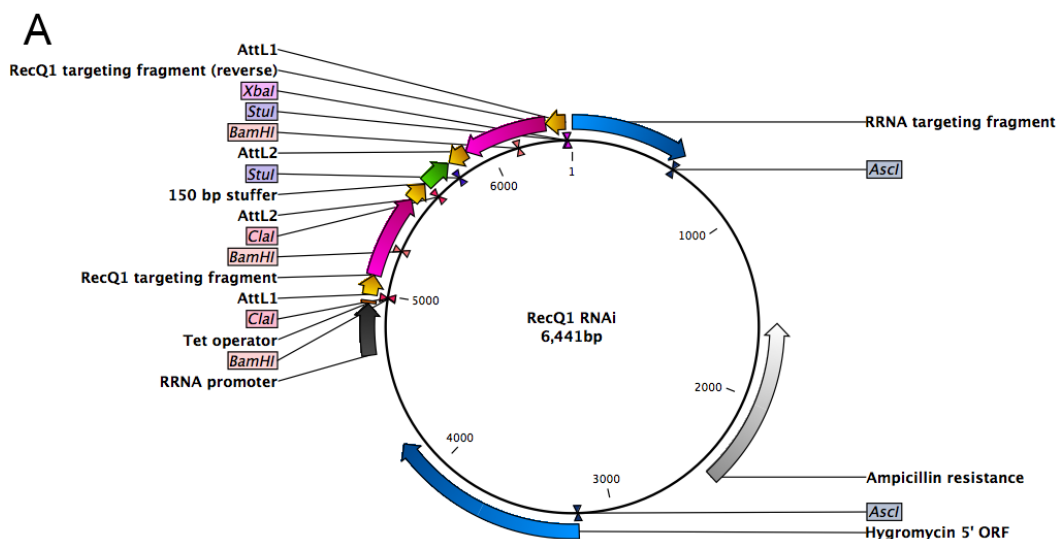


**Figure 5-5 Schematic of RNAi plasmid pGL2084**  
pGL2084 contains a the 5' region of the hygromycin (*HYG*) ORF and an *RRNA* targeting region (blue) for targeting integration of the construct into an *RRNA* spacer locus on chromosome 2a. Flanked by these targeting regions is an *RRNA* promoter, under the control of the Tet operator (dark grey), and two pairs of AttP sites (orange) for recombination of AttB-flanked PCR products. The pairs of AttP sites are linked by a 150 bp stuffer fragment (green), and flank a *ccdB* 'suicide gene' (light grey) for negative selection of improperly recombined plasmids after the Gateway reaction. Linearisation of the construct (after Gateway recombination) is by *Ascl*, which is indicated. Sizes are shown (bp).

The *RECQ1* RNAi targeting sequence was designed using RNAit (Redmond *et al.*, 2003). RNAit is an automated online tool for the selection of RNAi targets and primer sequences in *T. brucei*. Through NCBI BLAST analysis of the predicted PCR product against the *T. brucei* genome, RNAit is designed to exclude RNAi targeting fragments that target genes other than the gene of interest. A 425 bp *RECQ1* RNAi targeting sequence was designed (primers #67 and #68), which included the 29 nt AttB sequence at either end of the final PCR product. Primer sequences can be found in Appendix 7.1. A BP Gateway reaction was used to recombine two copies of the AttB-flanked *RECQ1* RNAi targeting fragment into pGL2084 in opposite directions (Fig 5-6A).

A schematic of a 2T1 cell line successfully transformed with a pGL2084-derived RNAi vector is shown in Figure 5-6B. The *EP* procyclin promoter drives expression of the *HYG* gene and the *RRNA* promoter drives transcription of the RNAi cassette, under the control of *TetO*. Absence of tetracycline represses transcription and addition of tetracycline induces transcription ( $2 \mu\text{g.mL}^{-1}$  in these experiments). Transcription of the two inverted *RECQ1* RNAi targeting

fragments and the stuffer generates a dsRNA structure connected by a 150 bp hairpin. This molecule should be recognised by the RNAi machinery and through the processes described above, target *RECQ1* mRNA for destruction, thus resulting in inhibition of RECQ1 protein synthesis.



**Figure 5-6 Schematic of RECQ1 RNAi construct, Gateway cloning and 2T1 transformation**

(A) The *RECQ1* RNAi plasmid was generated from Gateway recombination of *RECQ1* RNAi targeting fragments (pink) and the pGL2084 construct shown in Figure 5-5. Restriction sites used for confirmation of the *RECQ1* RNAi construct by restriction enzyme digest are indicated. Sizes are shown (bp). (B) The Gateway cloning process used to generate the *RECQ1* RNAi plasmid and its integration into 2T1 cells. A 425 bp region of the *RECQ1* ORF was amplified by PCR (primers #67 and #68), using primers that contained the 29 nt AttB sequence. The purified PCR product was recombined into the pGL2084 plasmid in a Gateway “BP clonase” reaction (Life Technologies), using the AttP recombination sites in pGL2084. The inverted orientation of the pair of AttP sites results in the integration of two PCR products in opposite orientations (indicated by black arrow). The pGL2084-derived *RECQ1* plasmid was digested with *Ascl* prior to transformation into 2T1 cells. Integration of the construct into the tagged *RRNA* spacer locus on chromosome 2a reconstitutes a complete hygromycin (*HYG*) ORF and removes the *PAC* gene (“puromycin”). Transcription of the *HYG* resistance gene is driven by an EP procyclin promoter (black arrow) and transcription of the RNAi cassette is driven by an *RRNA* promoter (black arrow), producing a dsRNA structure linked by the stuffer sequence (black box), which forms a hairpin. Figure reproduced from Catarina Marques, PhD Thesis (2015).

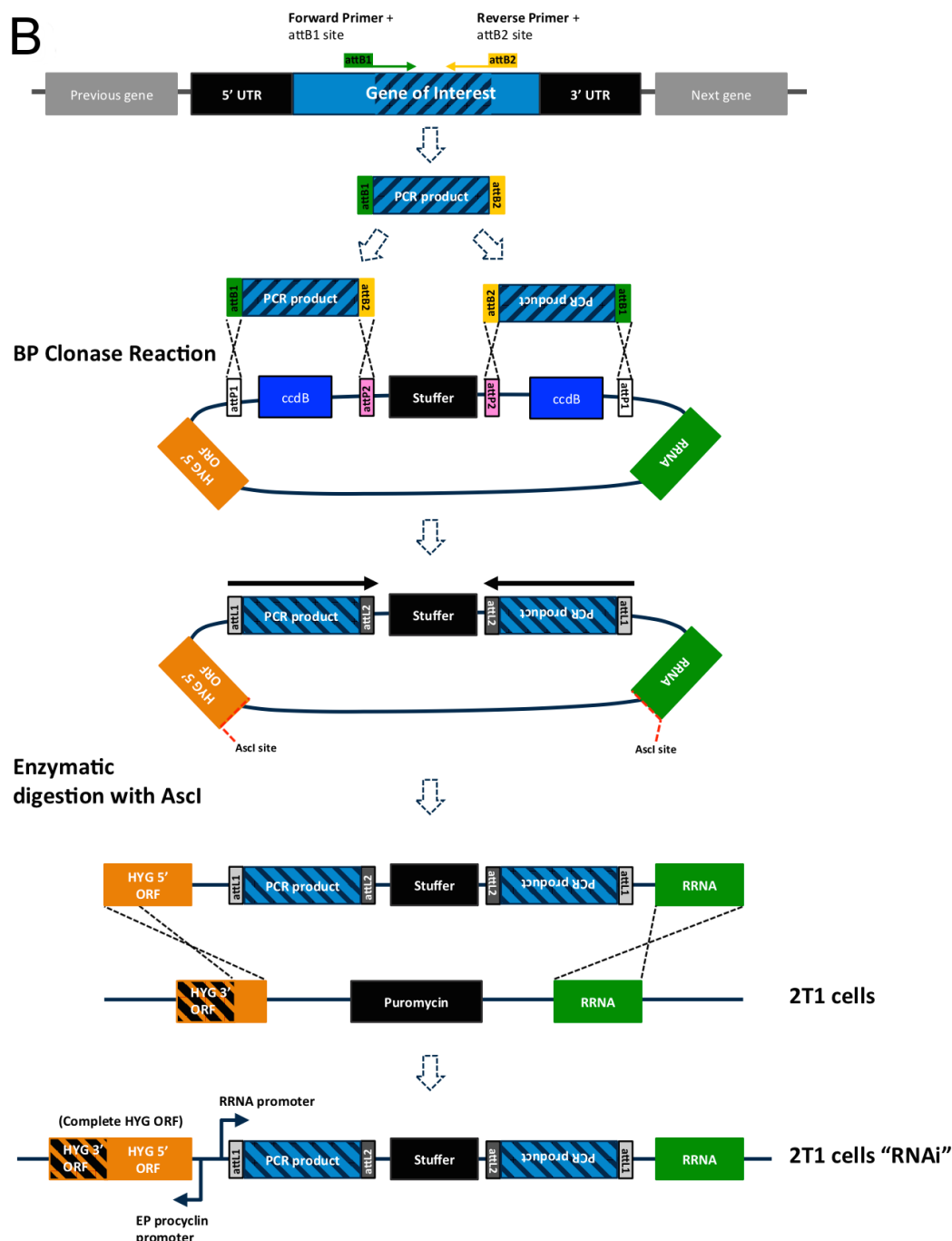
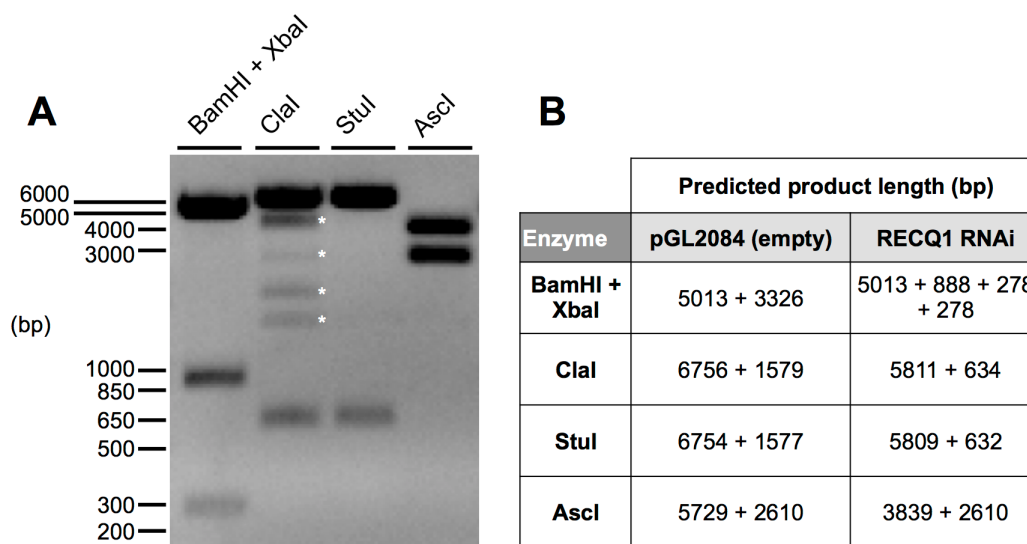


Figure 5-6 continued

Multiple restriction enzyme digests were used to confirm that the *RECQ1* RNAi construct (Fig. 5-6A) was correct and the results of these digests indicated incorrect recombination of the PCR products (data not shown). After multiple attempts, a correct plasmid was obtained. Figure 5-7 shows confirmation by restriction digest that the *RECQ1* RNAi plasmid was correct. One restriction enzyme digest, *Clal*, produced products in addition to the predicted ones. This was most likely due to the methylation sensitivity of *Clal*, or star activity in these conditions. As the *Clal* digest also produced the expected ~5.8 kb and



~600 bp digest products, and all of the other restriction enzyme digests produced the predicted products, the plasmid was considered to be correct and therefore used to transform 2T1 cells.



**Figure 5-7 Confirmation of RECQ1 RNAi plasmid**  
**(A)** The RECQ1 RNAi plasmid was tested for integration of the RNAi PCR product by restriction enzyme digestion using enzymes shown in **(B)**. White dots indicate digestion products in addition to those predicted. Sizes shown (ladder, bp). **(B)** Predicted sizes (bp) of products resulting from digestion of the empty pGL2084 vector and RECQ1 RNAi plasmid.

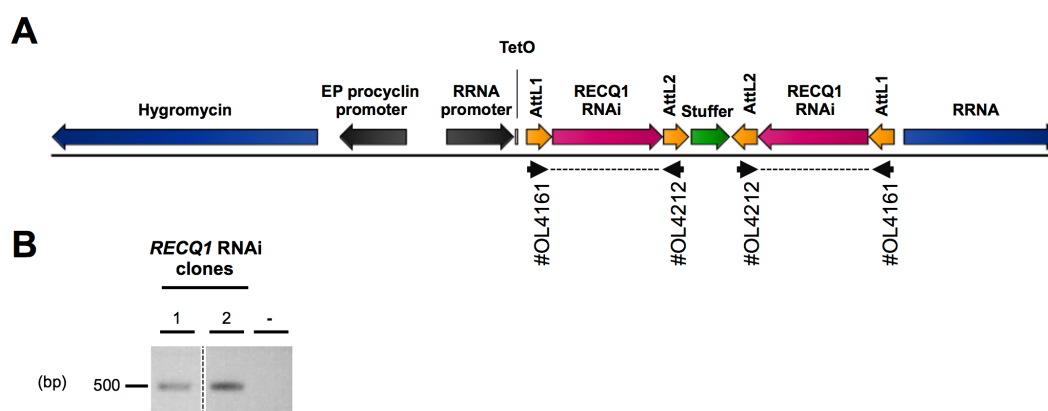
### 5.3.3 Generation of *RECQ1* RNAi cell line

2T1 cells were transformed with the AscI-linearised *RECQ1* RNAi construct and antibiotic resistant clones were selected with 2.5  $\mu\text{g}.\text{mL}^{-1}$  hygromycin and 0.5  $\mu\text{g}.\text{mL}^{-1}$  puromycin. Six putative *RECQ1* RNAi clones were obtained from the transformation. An additional 19 wells from the transformation contained live cells but as all 19 wells were on a single 24 well plate, they were unlikely to be clonal and therefore discarded. The puromycin sensitivity of the seven putative *RECQ1* RNAi clones was tested by assaying whether clones grew in the presence of 0.2  $\mu\text{g}.\text{mL}^{-1}$  puromycin. Clones in which the RNAi construct has integrated correctly are puromycin sensitive however, clones that are puromycin resistant but contain the RNAi construct can also be obtained (Alsford & Horn, 2008; Jones, 2013) and have presumably integrated the construct at a different locus. Four out of six putative *RECQ1* RNAi clones were puromycin sensitive and these were taken forward for phenotype analysis, as puromycin resistant clones have previously been demonstrated to give less consistent and stable RNAi phenotypes (Jones, 2013).

Preliminary analysis of *in vitro* growth of each of the four putative *RECQ1* RNAi clones was performed by induction of RNAi using 2  $\mu\text{g} \cdot \text{mL}^{-1}$  tetracycline. The cell density was then counted every 24 hours using a haemocytometer, up to a maximum of 120 hours. Cultures were diluted when appropriate to maintain a cell density of  $<1 \times 10^6 \text{ cells} \cdot \text{mL}^{-1}$  and the cumulative cell density was calculated. All four clones showed a very similar phenotype (discussed below, Section 5.4) and two were selected for more detailed analysis.

### 5.3.4 Confirmation of *RECQ1* RNAi cell line by PCR

The presence of the *RECQ1* RNAi cassette in the transformants was confirmed by PCR. Primers #OL4161 and #OL4212, which bind in the AttL sites formed by recombination of the *RECQ1* PCR product AttB and pGL2084 AttP sites, were used to amplify the *RECQ1* RNAi fragment (425 bp), as shown in Figure 5-8A. Figure 5-8B shows that both *RECQ1* RNAi clones used in experiments produced a PCR product of the expected size, suggesting they had integrated the *RECQ1* RNAi cassette.

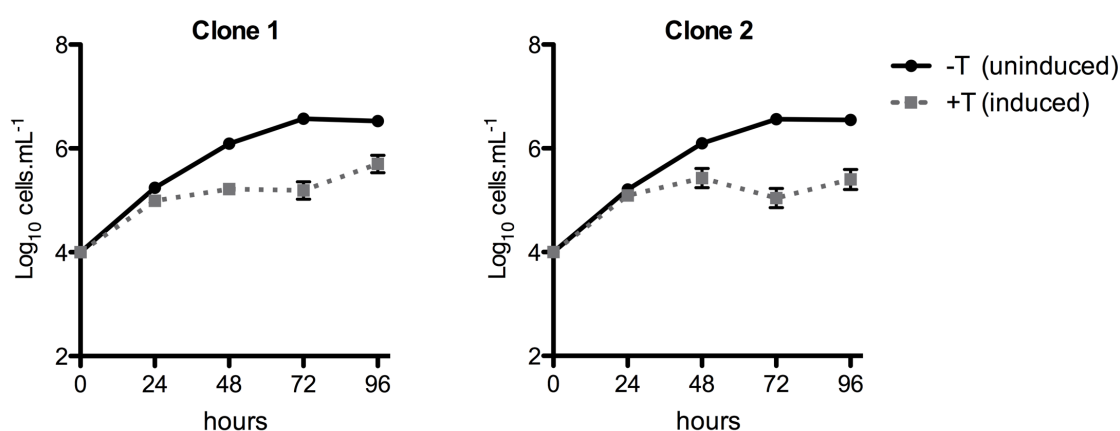


**Figure 5-8** PCR of *RECQ1* RNAi cassette integration in 2T1 cells  
(A) PCR primers #OL4161 and #OL4212 (black arrows under sequence) bind in the AttL sites of the RNAi cassettes (orange), amplifying the *RECQ1* RNAi targeting fragment (pink) to produce a 485 bp PCR product. (B) Genomic DNA from *RECQ1* RNAi Clone 1 and Clone 2 cells was analysed by PCR with primers #OL4161 and #OL4212 (see Appendix 7.1 for primer sequences). Distilled water was used as a negative control (-). Gaps indicate that lanes have been aligned in this figure after excision from multiple gels or from disparate parts of the same gel; size markers are shown (ladder, bp).

## 5.4 Phenotypic analysis of RECQ1 RNAi

### 5.4.1 Effect of RECQ1 RNAi knockdown on growth

To analyse the *in vitro* growth of BSF *T. brucei* following RECQ1 loss, RNAi was induced in cultures ( $2 \mu\text{g.mL}^{-1}$  tetracycline) at a density of  $1 \times 10^4 \text{ cells.mL}^{-1}$ . The cell density of cultures was counted every 24 hours up to 96 hours using a haemocytometer and cultures were not diluted during the time course. Growth curves were performed three times for both *RECQ1* RNAi clones. Induction of *RECQ1* RNAi resulted in a growth defect apparent from 48 hours post-induction, after which cessation of growth was seen in both *RECQ1* RNAi clones (Fig. 5-9); indeed, some evidence of cell death was seen in clone 2 from 48-72 hrs, as the cell density reduced. A slight increase in cell density was seen from 72-96 hours, which may result from the outgrowth of altered cells that no longer induce RNAi, though this was not examined in detail. These data appear consistent with the inability to generate a *recq1*<sup>-/-</sup> cell line (Section 5.3.1), and suggest that RECQ1 is essential for BSF growth.



**Figure 5-9** *In vitro* growth following *RECQ1* RNAi induction  
Mean cell counts of BSF *T. brucei* 2T1 cells following *RECQ1* RNAi induction by addition of  $2 \mu\text{g.mL}^{-1}$  tetracycline (broken line, +Tet) or without RNAi induction (solid line, -Tet) in two clones. Error bars indicate standard error of the mean (SEM).

### 5.4.2 Analysis of cell cycle

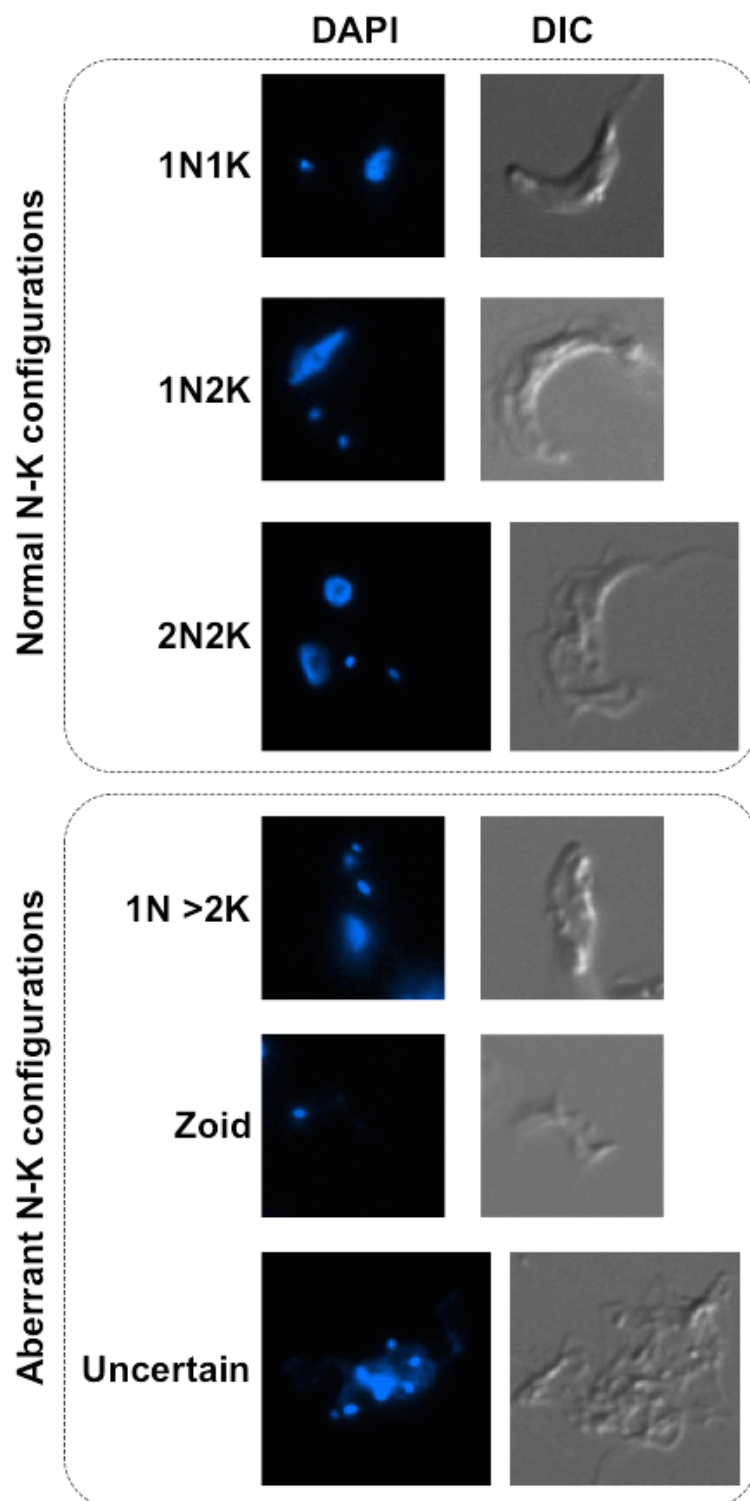
To investigate if the impaired growth following RNAi induction was associated with altered cell cycle distribution, quantification was performed by DAPI staining and analysing the proportion of 1N1K, 1N2K, 2N2K and abnormal cells by visualisation of the nucleus (N) and kinetoplast (K) (see Section 2.2.1). The number of nuclei and kinetoplasts correspond to different cell cycle stages in *T.*

*brucei*, providing a tool to examine progress through the cell cycle (see Section 3.4.2 for detailed explanation).

As mentioned previously (Section 4.6.1), the expected distribution of cell cycle phases in wild type BSF cells is: ~80% 1N1K, ~10% 1N2K, ~5% 2N2K (Benmerzouga *et al.*, 2013). In both clones, RNAi uninduced cells had a cell cycle distribution profile similar to this at all time points (Fig. 5-10A). In contrast, *RECQ1* RNAi induction severely altered this pattern. Within the first 48 hours post-RNAi induction, an increase in the proportion of 1N2K cells was seen in both clones, accompanied by a decrease in the 1N1K population. 1N2K cells represented ~30% of the total cell population 24 hours after induction, which suggests *RECQ1* knockdown results in a G<sub>2</sub>/M cell cycle arrest. However, at 48 hours post RNAi induction there was a high proportion (~40% in both clones) of aberrant 'other' cells. The proportion of 1N1K cells decreased further to ~30% at 48 hours post-induction, but the proportion of 1N2K cells remained similar (~20-25%) to 24 hours post-induction. At 48 hrs post-induction, the number of 2N2K cells appeared reduced, but the small number of such cells in the absence of RNAi makes this uncertain.

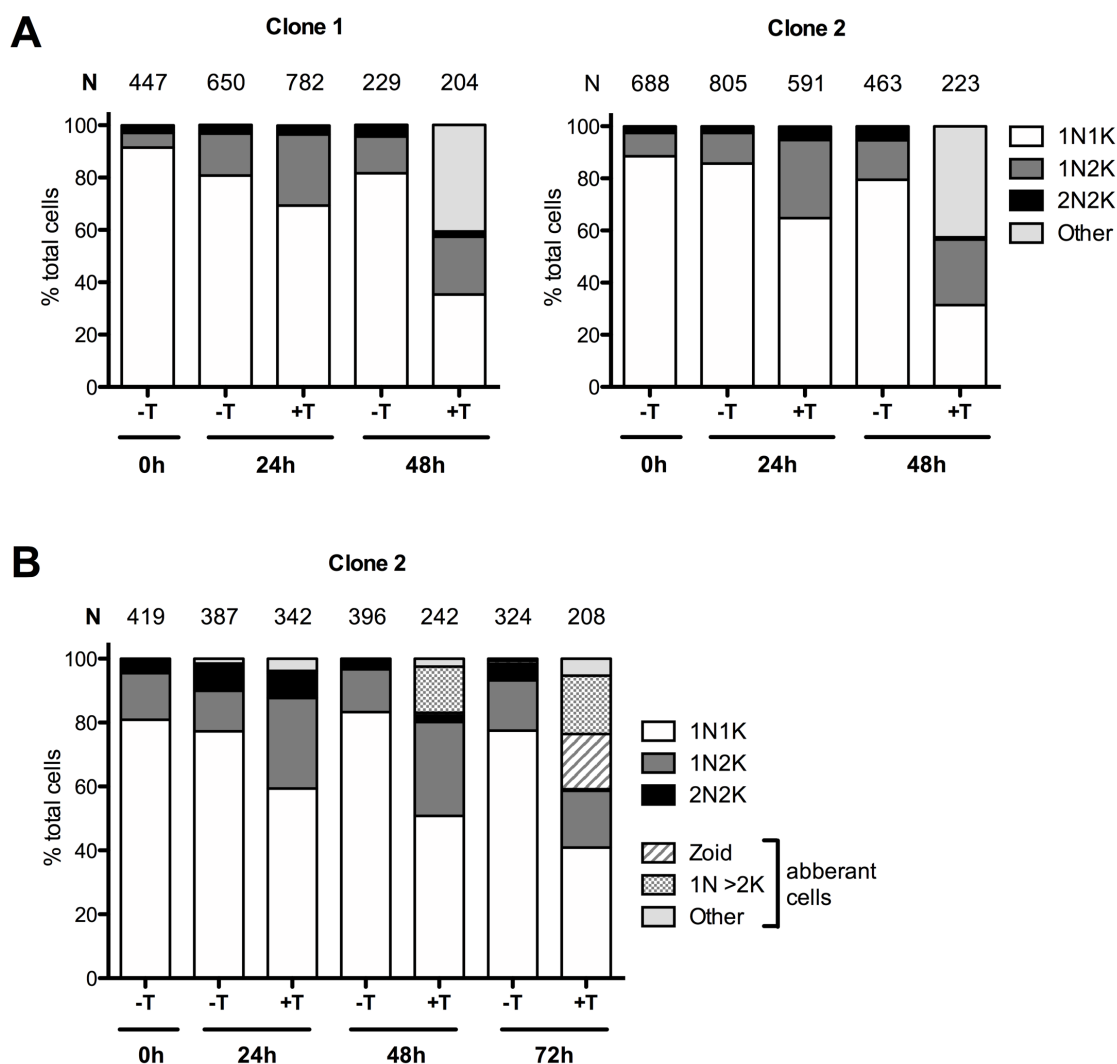
A number of different N-K configurations were observed at 48 hours post RNA-induction (Fig 5-10A). To examine these cell types, and the cell cycle perturbation after *RECQ1* RNAi further, the cell cycle analysis was extended to 72 hours in a further experiment using only *RECQ1* RNAi Clone 2 (chosen at random), sub-categorising the N-K configurations of 'other' cells (Fig. 5-11B). As observed in the first experiment, the proportion of 1N2K cells at 24 and 48 hours post-induction was similarly increased (~30% of the cell population, Fig. 5-11A), though appeared to reduce somewhat at 72 hours. Analysing the RNAi until 72 hours appeared to confirm the reduction in 2N2K cells, as these were barely detected at this time point after induction. At 48 hours post *RECQ1* RNAi induction, the proportion of aberrant cells was ~20% of the population (slightly lower than seen in the first analysis), and these were mainly comprised of 1N >2K cells, indicating perturbation of nuclear division. Extension of the time course to 72 hours of RNAi induction revealed the appearance of cells lacking a nucleus (0N1K, zoid cells), and an increase in 'others' relative to 48 hours (>40% of the population). The proportion of zoid and 1N >2K cells was approximately

equal at 72 hours (17% and 18% respectively), with a smaller proportion being difficult to classify.



**Figure 5-10** Cell types observed 48 hours post *RECQ1* RNAi induction  
 RNAi was induced in *RECQ1* RNAi [Clone 1 and Clone 2] cells by addition of tetracycline ( $2 \mu\text{g.mL}^{-1}$ ) and cell samples taken at 0, 24 and 48 hours post-induction. Cells were stained with DAPI to visualise the nucleus and the number of each cell type. Examples are shown of normal (1N1K, 1N2K and 2N2K) and aberrant cell types observed after 48 hours of *RECQ1* RNAi induction. N, nucleus; K, kinetoplast.

These data suggest that RECQ1 may play a role in nuclear DNA replication. Knockdown of RECQ1 results in an accumulation of 1N2K cells, indicating cell G2/M cycle arrest followed by the appearance of aberrant cells, the vast majority of which are zoids or 1N >2K. Taken together, these observations suggest that RECQ1 RNAi induces a checkpoint due to problems with nuclear replication or division, but that this is not absolute and cells emerge in which the nuclei is lost (zoids) or in which the kinetoplast continues to replicate without associated nuclear replication and division (1N >2K). Further analysis of the DNA content of cells by FACS following RECQ1 knockdown would be necessary to confirm this hypothesis, as would a directed experiment to assess DNA synthesis, such as measuring nucleotide analogue incorporation.



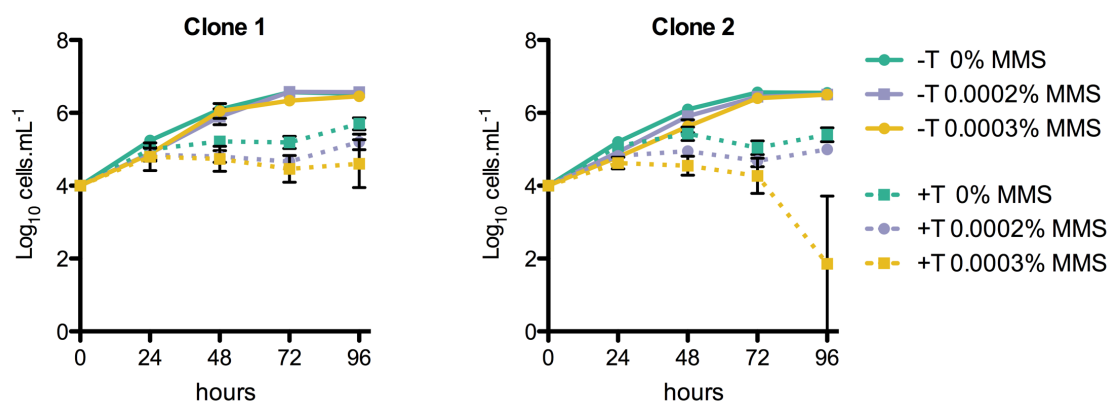
**Figure 5-11** Cell cycle analysis following *RECQ1* RNAi induction  
(A) RNAi was induced in *RECQ1* RNAi [Clone 1 and Clone 2] cells by addition of tetracycline ( $2 \mu\text{g} \cdot \text{mL}^{-1}$ ) and cell samples taken at 0, 24 and 48 hours post-induction. Cells were stained with DAPI to visualise the nucleus and the number of each cell type (1N1K, 1N2K, 2N2K & other) was counted and is represented as a percentage of cells counted. The number of cells counted is indicated (N). (B) *RECQ1* RNAi induction and DAPI analysis was repeated in Clone 2, additionally analysing cells at 72 hours post induction. Aberrant cells were classified as 1N >2K, zoid, and other. The number of cells counted is indicated (N).

### 5.4.3 Analysis of methyl methanesulfonate damage sensitivity after RNAi

To investigate whether *RECQ1* plays a role in DNA repair, the sensitivity of BSF *T. brucei* to the DNA damaging agent MMS following *RECQ1* RNAi knockdown was analysed. This was performed by measurement of *in vitro* growth following RNAi induction in the presence of 0.0002% and 0.0003% MMS and compared with growth of uninduced cells after MMS treatment.

*RECQ1* RNAi was induced in 1 mL cultures at a density of  $1 \times 10^4$  cells.mL<sup>-1</sup> by addition of  $2 \mu\text{g} \cdot \text{mL}^{-1}$  tetracycline and MMS was added to the appropriate final

concentrations. The cell density was counted using a haemocytometer every 24 hours up to 96 hours without dilution of the culture. All growth curves were repeated three times. MMS appeared to have little effect on the growth of uninduced cells (Fig. 5-12). Induction of RNAi again arrested the growth of both clones and, in these conditions, the addition of MMS at either concentration further impaired the cells' growth, an effect that was most marked in Clone 2 at 0.0003% MMS at 96 hours, where most cells were dead and therefore the cell density was too low to measure using a haemocytometer. From these data it would seem that RNAi knockdown of RECQ1 results in slightly increased MMS sensitivity of BSF *T. brucei*, though whether this represents evidence for a role in damage repair or exaggeration of the growth defect caused by RNAi is unclear.



**Figure 5-12 Cell cycle analysis following *RECQ1* RNAi induction**  
 RNAi was induced in *RECQ1* RNAi [Clone 1 and Clone 2] cells by addition of tetracycline ( $2 \mu\text{g.mL}^{-1}$ ) at a cell density of  $1 \times 10^4 \text{ cells.mL}^{-1}$  and MMS was added to a final concentration of 0%, 0.0003% or 0.0004%. The cell density was measured using a haemocytometer every 24 hours up to 96 hours. The mean cell density of three independent cultures are shown, with bars representing standard error of the mean (SEM).

#### 5.4.4 Analysis of $\gamma\text{H2A}$ intensity following *RECQ1* RNAi

As discussed in Chapter 4 (see Section 4.6.3) phosphorylated variant histone H2A ( $\gamma\text{H2A}$ ) is an early marker of DNA damage in eukaryotes (Canman, 2003; Polo & Jackson, 2011). It is also detectable in BSF *T. brucei* following treatment with DNA damaging agents or induction of DSBs (Glover *et al.*, 2013a; Glover & Horn, 2012). To investigate whether knockdown of RECQ1 results in an increased level of endogenous DNA damage, the  $\gamma\text{H2A}$  level in cells was quantified by immunolocalisation of  $\gamma\text{H2A}$  and measurement of nuclear fluorescence intensity.

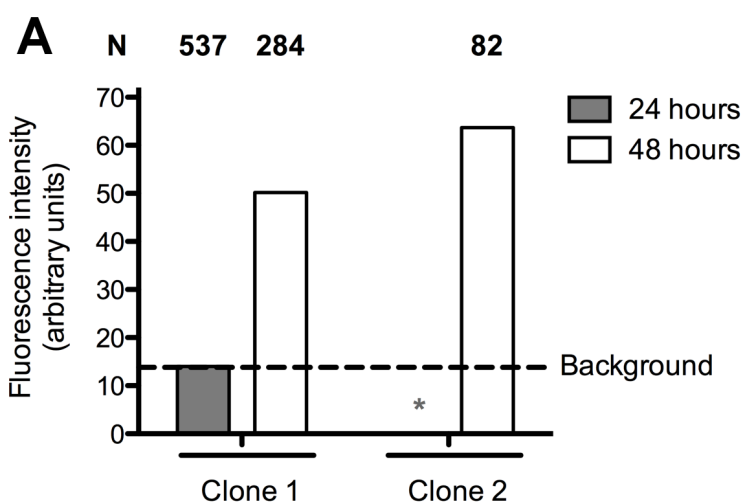


RNAi was induced in cultures at  $5 \times 10^4$  cells.mL<sup>-1</sup> by addition of 2 µg.mL<sup>-1</sup> tetracycline and cell samples were taken at 0, 24 and 48 hours. A higher starting cell density was used in order to collect enough RNAi induced cells for analysis at 48 hours. Cells were fixed on glass slides and stained for γH2A using rabbit anti-γH2A antiserum (gift; Tiago D Serafim, unpublished; 1:1000 dilution) and secondary Alexa Fluor 594-conjugated anti-rabbit antiserum (Life Technologies, 1:2000 dilution). All slides were stained at the same time using the same antisera working solutions. The intensity of γH2A signal in the nucleus was calculated by measuring the average intensity of all nuclei in captured images using ImageJ ([imagej.nih.gov/ij](http://imagej.nih.gov/ij)) (Rasband, 1997-2014), and is presented in arbitrary units. The background level of γH2A fluorescence was calculated as the average γH2A nuclear intensity in uninduced cells for both clones and all time points.

In Clone 1 the nuclear γH2A fluorescence of cells at 24 hours post *RECQ1* RNAi induction was unchanged from the background level of nuclear γH2A fluorescence in the uninduced cells (Fig 5-13). At the 24 hour time point it was not possible to analyse RNAi induced Clone 2 cells due to malfunctions in capture of these images. Nevertheless, at 48 hours post RNAi induction, the nuclear γH2A intensity of both RNAi induced cells was four to five-fold higher than background levels in both clones, indicating an increase in γH2A in the nucleus. This suggests that knockdown of *RECQ1* results in an increased level of DNA damage apparent at 48 hours, the same time point at which a pronounced growth defect is first observed (Fig 5-9).

γH2A expression has previously been shown to increase following knockdown of a RecQ protein. Knockdown of the human *RECQ1* protein, which renders cells more sensitive to DNA damage, results in spontaneous γH2A formation (Sharma & Brosh, 2007). As previously discussed, phosphorylation of variant histone H2A occurs following DNA damage. DNA damage can be incurred from multiple sources, exogenous and endogenous. RecQ helicases play multiple roles in maintaining genome integrity, including replication fork integrity and recovery, and in repair of DNA damage. If *RECQ1* plays a role in replication, for example maintaining recovering stalled forks to prevent DNA lesions as some other RecQ helicases do (Berti *et al.*, 2013; Zeman & Cimprich, 2014), knockdown of *RECQ1* would result in an increase in DNA lesions from collapsed replication forks,

triggering H2AX phosphorylation. RecQ helicases also function in the repair of non-replicative damage, such as DSB repair by the yeast SGS1 RecQ helicase (Mimitou & Symington, 2008). A similar role for RECQ1 would also lead to an increase in  $\gamma$ H2A expression. Analysis of RECQ1 expression through the cell cycle would be useful in investigating whether RECQ1 plays a role in nuclear DNA replication or the repair of replicative damage. Cell cycle dependent expression is not necessary for, but would be consistent with, a role in either of these processes.



**Figure 5-13** Immunofluorescent analysis of nuclear  $\gamma$ H2A intensity following *RECQ1* RNAi induction

RNAi was induced in *RECQ1* RNAi Clone 1 and Clone 2 cells at  $5 \times 10^4$  cells.mL<sup>-1</sup> (2  $\mu$ g.mL<sup>-1</sup> tetracycline). Samples were taken at 0, 24 and 48 hours post induction, fixed, stained for  $\gamma$ H2A using rabbit anti- $\gamma$ H2A antiserum (gift; Tiago D Serafim, unpublished; 1:1000 dilution) and secondary Alexa Fluor 594-conjugated anti-rabbit antiserum (Life Technologies, 1:2000) and DNA visualised using DAPI. The  $\gamma$ H2A fluorescence intensity is represented as the average nuclear  $\gamma$ H2A fluorescence of cells counted (N, number of cells counted), calculated using ImageJ (arbitrary units). Background fluorescence is the average  $\gamma$ H2A nuclear intensity of all uninduced cells. No cells were counted at the 24 hour time point for Clone 2 (\*).

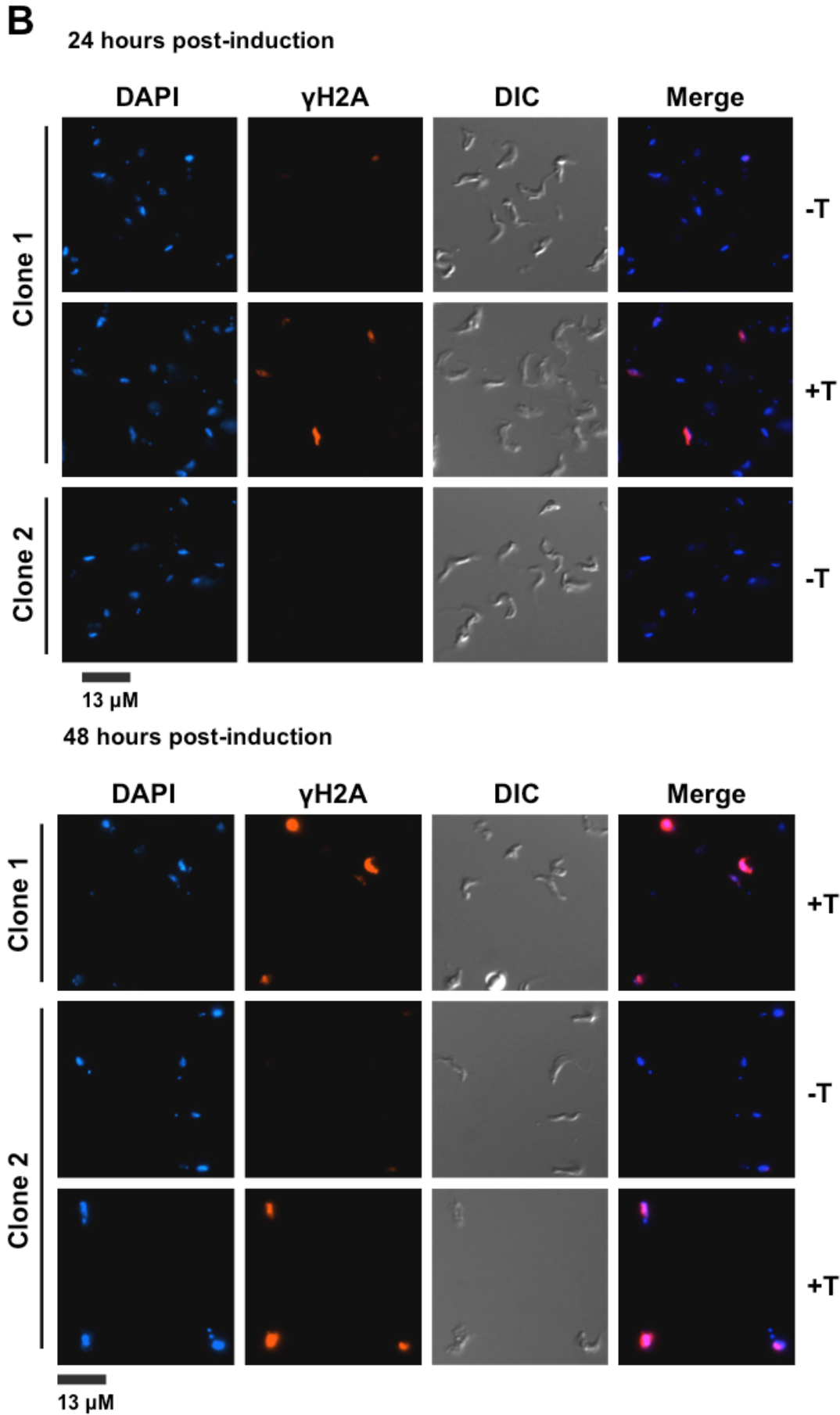
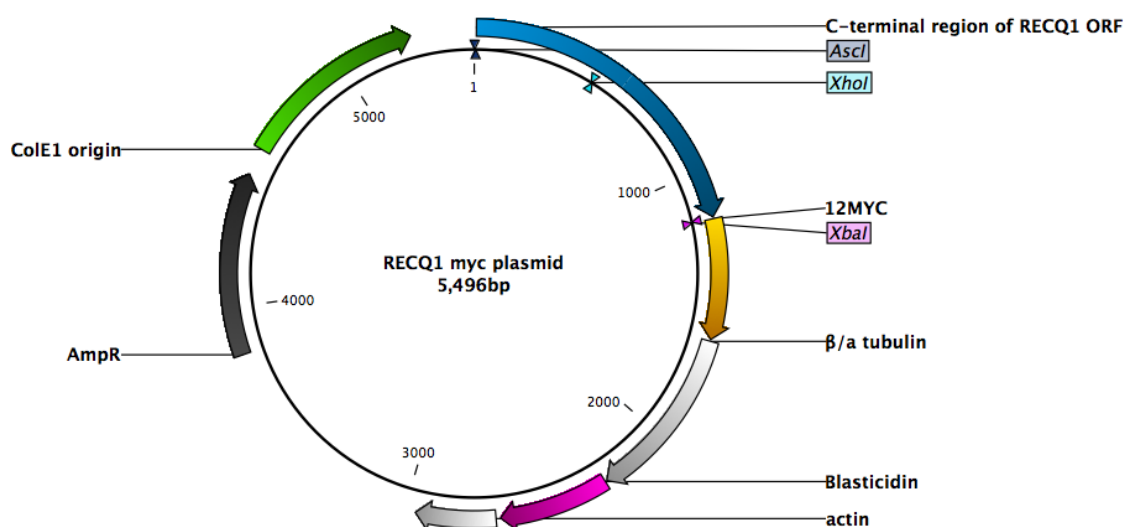


Figure 5-13 continued

### 5.4.5 Confirmation of RECQ1 knockdown by western blot analysis

In order to confirm knockdown of RECQ1 in both *RECQ1* RNAi clones, one allele of RECQ1 was endogenously tagged with a 12myc epitope. The same backbone construct (pNAT<sup>x12myc</sup>) as used to epitope tag MUS81 and PIF6 was used to tag RECQ1 (see Section 3.7 for detailed explanation). The construct (Fig. 5-14) contains the final 1.1 kb of the *RECQ1* coding sequence, excluding the STOP codon, cloned upstream of and in frame with a 12myc coding sequence. The cloned *RECQ1* region contained a unique restriction enzyme site (*Xho*I) for linearisation of the construct prior to transformation. The *RECQ1* region was cloned into the construct using *Asc*I (5') and *Xba*I (3') restriction enzyme sites, and confirmed by restriction digestion and sequencing. Selection of transformants is enabled by a blasticidin resistance gene, flanked by tubulin and actin processing sequences. The blasticidin cassette provides polyadenylation signals for RECQ1 12myc, making the 3' UTR of RECQ1 12myc non-endogenous.



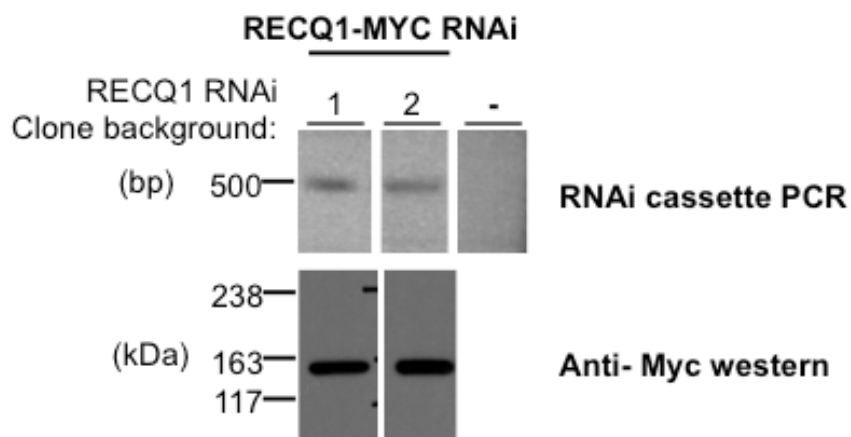
**Figure 5-14** Schematic of the *RECQ1* 12myc plasmid  
The *RECQ1* 12myc plasmid contains 1.1 kb of the 3' *RECQ1* ORF up to but not including the stop codon (orange), followed by 12 copies of the myc epitope sequence (blue). The unique restriction enzyme site (*Xho*I) used to linearise the construct is indicated. The blasticidin resistance cassette is shown (pink) flanked by  $\beta$ a tubulin and actin intergenic sequences (light grey). Ampicillin resistance gene (AmpR, dark grey) and the ColE1 origin (green) are also shown. Sizes shown (bp).

Both *RECQ1* RNAi clones were transformed with the RECQ1 12myc construct, linearised using *Xho*I, and antibiotic resistant transformants were selected using 10  $\mu\text{g} \cdot \text{mL}^{-1}$  blasticidin. Blasticidin-resistant clones were screened by western blot and PCR to check that they expressed myc-tagged RECQ1 and retained the *RECQ1* RNAi cassette (Fig 5-15). Whole cell lysates were separated by SDS-PAGE

on a 3-8% tris-acetate gel, western blotted and probed with mouse anti-myc antiserum (Millipore, 1:7000) and secondary anti-rabbit HRP-conjugated antiserum (Life Technologies, 1:5000). The presence of the RNAi cassette was confirmed by PCR amplification of a 485 bp region of the cassette (#OL4161 and #OL4212, sequences in Appendix 7.1). All putative RECQ1-myc RNAi clones retained the RNAi cassette, but only seven out of 10 RECQ1-12myc RNAi clones in the RECQ1 RNAi Clone 1 and eight out of twelve clones in the RECQ1 RNAi Clone 2 background expressed a protein of the expected size of RECQ1 12myc (116.2 kDa). Figure 5-15 shows the western blot and PCR analysis of the RECQ1-myc RNAi clones used in the experiments described below.

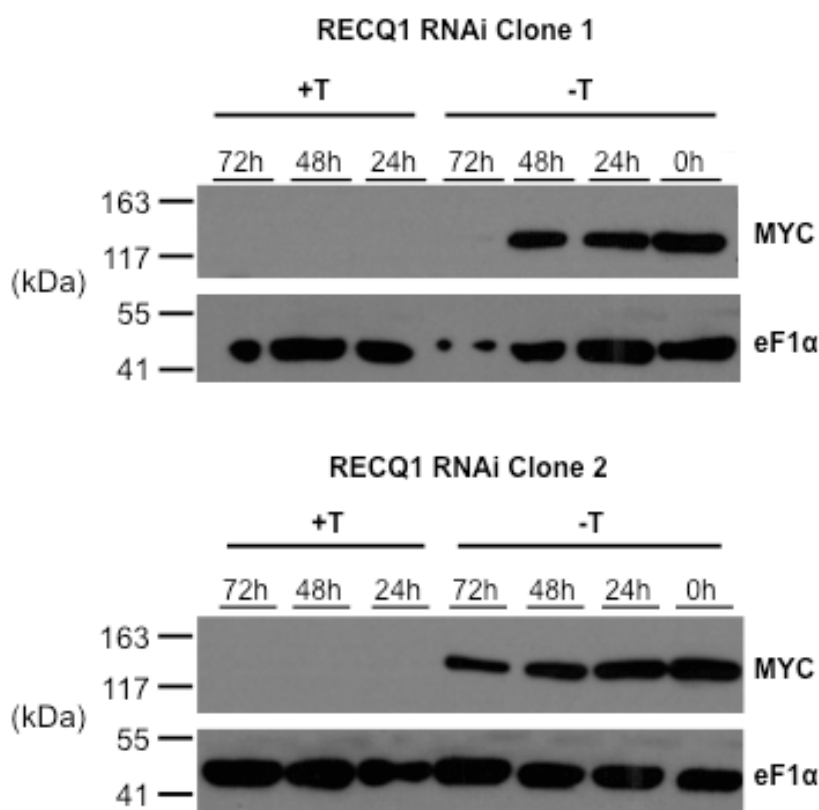
To test for knockdown of RECQ1 at the protein level, RECQ1 12myc expression following RNAi induction in RECQ1myc RNAi was assayed by western blot. RNAi was induced in both RECQ1myc RNAi clones by addition of tetracycline ( $2 \mu\text{g} \cdot \text{mL}^{-1}$ ) to cultures at a density of  $1 \times 10^5 \text{ cells} \cdot \text{mL}^{-1}$ . Cells were collected at 0, 24, 48 and 72 hours after induction or in non-induced controls. A higher starting cell density was used to enable collection of enough cells for western blot analysis in RNAi induced cultures at later time points. RNAi uninduced cultures were diluted to  $1 \times 10^5 \text{ cells} \cdot \text{mL}^{-1}$  at 24 and 48 hours to prevent cultures entering stationary phase. At these time points, RNAi induced cultures were resuspended in fresh medium to the cell density measured at that time point and fresh tetracycline added (final  $2 \mu\text{g} \cdot \text{mL}^{-1}$ ). Whole cell lysates were separated by SDS-PAGE on a 3-8% tris-acetate gel, with the same number of cell equivalents loaded in each lane. Gels were western blotted and probed with mouse anti-myc antiserum and secondary anti-mouse HRP-conjugated antiserum to detect RECQ1 12myc, and mouse anti-EF1 $\alpha$  and anti-mouse HRP-conjugated antisera as a loading control.

Figure 5-16 shows RECQ1 12myc is not detectable in the RECQ1 RNAi Clone 1 or the RECQ1 RNAi Clone 2 background at 24 hours post RNAi-induction, or at any point thereafter. Similar expression of EF1 $\alpha$  in all samples confirms similar protein loading in all samples. Both RECQ1 RNAi Clone 1 and Clone 2, therefore, exhibited a strong knockdown of RECQ1 protein upon induction by the addition of tetracycline.



**Figure 5-15** Western blot confirmation of RECQ1-12myc expression in RECQ1 12myc RNAi clones

Genomic DNA from RECQ1-myc RNAi clones was assayed by PCR for the presence of the RNAi cassette (primers #OL4161 and #OL4212; 485 bp PCR product). Dashed lines indicate that lanes have been aligned in this figure after excision from multiple gels or from disparate parts of the same gel; size markers are shown (ladder, bp). Total protein extract from RECQ1 12myc RNAi cells was separated by SDS-PAGE on a 3-8% tris-acetate gel, western blotted and probed with mouse anti-myc antiserum (Millipore, 1:7000 dilution) and secondary anti-rabbit HRP-conjugated antiserum (Life Technologies, 1:5000 dilution). Predicted size of RECQ1 12myc (116.2 kDa; RECQ1 101.8 kDa + 12myc 14.4 kDa). Size markers are shown (kDa).



**Figure 5-16** Confirmation of *RECQ1* knockdown by western blot analysis

Whole cell lysates ( $5 \times 10^6$  cell equivalents) of RNAi induced (+T) and uninduced (-T) cultures were separated by SDS-PAGE on a 3-8% tris-acetate gel, western blotted and probed with mouse anti-myc antiserum (Millipore, 1:7000 dilution) and secondary anti-mouse HRP-conjugated antiserum (Life Technologies, 1:5000) to detect RECQ1 12myc, and mouse anti-eF1α antiserum (Millipore, 1:25,000 dilution) and anti-mouse HRP-conjugated antiserum (Life Technologies, 1:5000 dilution) as a loading control. -T, no tetracycline

added (control); +T, tetracycline added to 2  $\mu\text{g.mL}^{-1}$ ; clone 1, *RECQ1* RNAi clone 1 background; clone 2, *RECQ1* RNAi clone 2 background. Size markers are shown (kDa).

## 5.5 Immunolocalisation of RECQ1

Immunolocalisation of RECQ1 and co-immunolocalisation with RAD51 was carried out following DNA damage in order to investigate the cellular localisation of RECQ1 before and after the induction of DNA damage and test whether it co-localises with RAD51. This was achieved using endogenously C-terminal 12myc epitope tagged RECQ1 and anti-RAD51 antiserum. This strategy has previously been successfully used to tag and localise proteins in *T. brucei* (Trenaman *et al.*, 2013). As discussed previously, RAD51 catalyses DNA strand exchange during DNA repair and *T. brucei* cells in which DNA damage has been induced form subnuclear RAD51 foci (Hartley & McCulloch, 2008; Proudfoot & McCulloch, 2005). The DNA repair factor BRCA2 has been demonstrated to co-localise with these RAD51 foci in a limited form (Trenaman *et al.*, 2013). RECQ2 also co-localises with RAD51 repair foci, as discussed in Chapter 3 (see Section 3.9.1). Co-localisation of RECQ1 with RAD51 foci following DNA damage would suggest that RECQ1 plays a role in the repair of DNA damage.

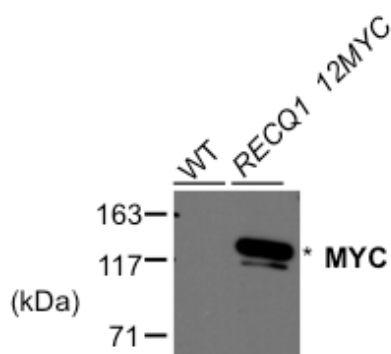
### 5.5.1 Generation of an endogenously myc-tagged RECQ1 cell line

The RECQ1 12myc construct used previously to myc-tag RECQ1 in *RECQ1* RNAi cell lines (see Section 5.4.5) was used to myc-tag RECQ1 in wild type Lister 427 cells. See Section 5.4.5 for details of the RECQ1 12myc construct.

### 5.5.2 Confirmation of myc-tagged line by western blot analysis

Wild type BSF Lister 427 cells were transformed with the linearised RECQ1 12myc construct in order to generate a C-terminal 12myc tagged RECQ1 cell line. Antibiotic resistant transformants were selected using 10  $\mu\text{g.mL}^{-1}$  blasticidin. Antibiotic resistant clones obtained from this transformation were screened by western blot performed on total protein extracts. Total protein extracts were separated by SDS-PAGE on a 3-8% tris-acetate gel before western blotting and probing with mouse anti-myc antiserum (Millipore, 1:7000 dilution) and HRP-conjugated anti-mouse secondary antiserum (Life Technologies, 1:5000 dilution). All of the four putative clones tested expressed a protein of the expected size of RECQ1 12myc (116.2 kDa; RECQ1 101.8 kDa + 12myc 14.4 kDa).

Western blot analysis of the clone taken forward for use in all experiments described below is shown in Figure 5-17.



**Figure 5-17** Confirmation of RECQ1 12myc expression by western blot  
Total protein extract from putative RECQ1 12myc cells was separated by SDS-PAGE on a 3-8% tris-acetate gel, western blotted and probed with mouse anti-myc antiserum (Millipore, 1:7000 dilution) and secondary anti-mouse HRP-conjugated antiserum (Life Technologies, 1:5000 dilution). A band of the size expected for RECQ1 12myc (116.2 kDa; RECQ1 101.8 kDa + 12myc 14.4 kDa) is indicated (\*). Size markers are shown (kDa).

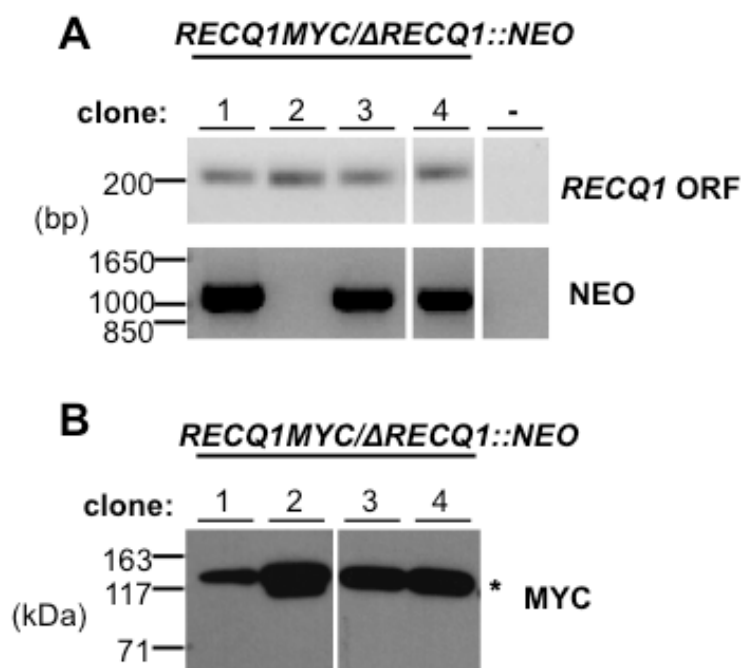
### 5.5.3 Testing the functionality of endogenously-tagged RECQ1

To test whether RECQ1 12myc is functional, the second (untagged) *RECQ1* allele was knocked out, replacing the ORF with a *NEO* resistance cassette, as performed when testing the functionality of RECQ2myc, MUS81myc and PIF6myc in Chapter 3 (see Section 3.8). As described above (Section 5.3), *RECQ1* is essential. Therefore, if RECQ1 12myc is functional, successful generation of *RECQ1 12myc/ΔRECQ1::NEO* cells should be possible. In contrast, if it is non-functional, attempts at generating *RECQ1 12myc/ΔRECQ1::NEO* cells should fail.

The construct used to knockout the untagged *RECQ1* allele, *ΔRECQ1::NEO*, was the same construct as that used in the attempt to generate *recq1*<sup>-/-</sup> mutants (Section 5.3.1). RECQ1 12myc cells were transformed with the linearised *ΔRECQ1::NEO* construct. Antibiotic resistant cells were selected with 5 μg.mL<sup>-1</sup> blasticidin and 2.5 μg.mL<sup>-1</sup> G418. Four putative *RECQ1 12myc/ΔRECQ1::NEO* clones were screened by PCR for integration of the *ΔRECQ1::NEO* construct (“NEO”, primers #50 and #155) using the same PCR strategy as described in Section 3.3.3 and a ~200 bp region of the *RECQ1* ORF was amplified as a control (“*RECQ1* ORF”, primers #75 and #76). Clones were additionally screened by western blot analysis, where whole cell lysates were separated by SDS-PAGE on a 3-8% tris-acetate gel, western blotted and probed with mouse anti-myc antiserum (Millipore, 1:7000 dilution) followed by anti-mouse HRP-conjugated



antisera (Life Technologies, 1:5000 dilution). PCR and western blot analysis showed that three out of four putative *RECQ1* 12myc/ $\Delta$ *RECQ1*::*NEO* clones had integrated the knockout cassette at the correct locus and still expressed *RECQ1* 12myc (Fig. 5-18). As *RECQ1* is essential, the successful generation of *RECQ1* 12myc/ $\Delta$ *RECQ1*::*NEO* cells suggests that *RECQ1* 12myc is functional. However, it is possible that an ‘extra’ wild type *RECQ1* allele has been formed following integration of the *NEO* knockout construct, as the PCR tests outlined above did not test for this.



**Figure 5-18** Confirmation of successful *RECQ1* 12myc/ $\Delta$ *RECQ1* by PCR and western blot analysis

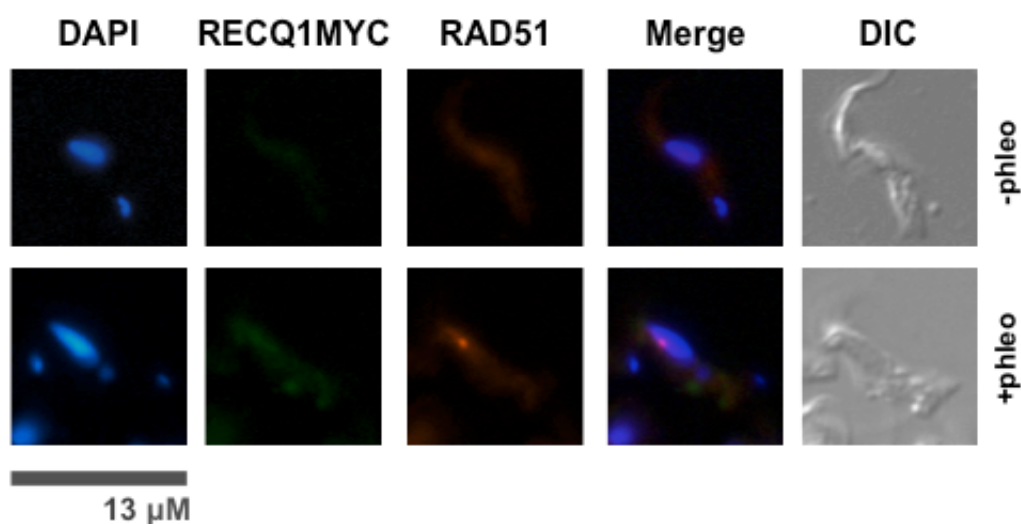
(A) PCR amplification of a region of *RECQ1* (226 bp, primers #75 and #76) and PCR amplification using primers #50 & #155 (1085 bp) that bind in the *NEO* resistance cassette and upstream of *RECQ1*. Primer sequences can be found in Appendix 7.1. Sizes are shown (bp). (B) Total protein extracts were separated by SDS-PAGE, western blotted and probed with mouse anti-myc antiserum (Millipore, 1:7000 dilution) and anti-mouse HRP-conjugated antisera (Life Technologies, 1:5000 dilution). A band of the expected size of *RECQ1* 12myc is indicated by (\*). Size markers are shown (kDa).

#### 5.5.4 Localisation of *RECQ1*myc

Immunolocalisation of *RECQ1* 12myc and co-localisation with RAD51 was carried out as previously described (see Sections 3.9 and 2.10). *RECQ1* 12myc cultures ( $1 \times 10^6$  cells.mL<sup>-1</sup>) were incubated for 18 hours in the presence or absence of 1  $\mu$ g.mL<sup>-1</sup> phleomycin. Cells were fixed and stained with anti-myc Alexa Fluor 488-conjugated antiserum (Millipore, 1:2000 dilution) and co-stained with rabbit anti-RAD51 antiserum (Proudfoot & McCulloch, 2005) (1:500 dilution) and anti-

rabbit Alexa Fluor 594 conjugated antiserum (Life Technologies, 1:7000 dilution). DNA was stained with DAPI for visualisation of the kinetoplast and nucleus.

As shown in Figure 5-19, RECQ1 12myc signal was not observed in any cells either in the presence or absence of phleomycin. In contrast, in these conditions, visible RAD51 subnuclear foci were seen, as described previously. The absence of detectable subcellular RECQ1 12myc signal is not due to absence of RECQ1 12myc expression, as western analysis confirms that RECQ1 12myc is expressed in this cell line (see Figure 5-16, Section 5.5.2). Nor is it a problem with the staining procedure, as slides to localise 12myc RECQ2, MUS81 12myc and PIF6 12myc were prepared using the same solutions and signal was observed in these (see Figures 3-30 to 3-32, Section 3.9).



**Figure 5-19** Representative examples of attempts at RECQ1 12myc immunolocalisation. RECQ1 12myc cells were incubated for 18 hours in the presence or absence of phleomycin ( $1 \mu\text{g} \cdot \text{mL}^{-1}$ ). Alexa Fluor 488 antiserum was used to stain RECQ1 12myc (RECQ1myc), RAD51 was localised using anti-RAD51 antiserum (RAD51) and goat anti-rabbit Alexa Fluor 594 conjugated antiserum, DAPI was used to visualise the DNA (DAPI) as described in the text. Differential interference contrast was used to visualise whole cells (DIC). Merged DAPI and RAD51 images are also shown (merge). Scale,  $13 \mu\text{M}$ .

The absence of RECQ1 12myc signal could be due to low expression of RECQ1 12myc to a level not detectable by this method. The absence of RECQ1 12myc foci following phleomycin damage, unlike RAD51 and BRCA2 (Proudfoot & McCulloch, 2005; Trenaman *et al.*, 2013), suggests that RECQ1 12myc is not involved in the response to phleomycin-induced DNA damage. In comparison with the strong cell cycle defects after *RECQ1* RNAi, the lack of detectable subnuclear localisation after phleomycin treatment and the

limited sensitivity to MMS suggest that any role for RECQ1 in repair is subservient to a function in genome duplication.

## 5.6 Summary

The aim of this chapter was to investigate the function of TbRECQ1, in particular to ask whether it plays a role in DNA repair. This was initially attempted using a gene knockout approach but, due to an inability to obtain a *recq1*<sup>-/-</sup> mutant, an RNAi approach was adopted. Analysis of the growth, cell cycle profile, MMS sensitivity and γH2A expression following *TbRECQ1* RNAi induction was undertaken. Additionally, localisation of TbRECQ1 was attempted using endogenous myc tagging, which was successful but did not reveal any subcellular signal using immunofluorescence. To conclude this chapter, the data from analysis of the phenotype of BSF *T. brucei* following RECQ1 knockdown and hypotheses for the possible function(s) of TbRECQ1 will be discussed.

### 5.6.1 TbRECQ1 is a diverged RecQ-like helicase

Analysis of the predicted protein sequence of TbRECQ1 showed that it is predicted to contain some but not all of the protein domains usually found in RecQ-like helicases. TbRECQ1 lacks both an HRDC domain, usually found in RecQ-like helicases. Phylogenetic analysis of *TbRECQ1* and a number of other bacterial and eukaryotic RecQ-like helicases was undertaken. This showed that TbRECQ1, along with some other kinetoplastid RecQ-like proteins, groups more distantly from most RecQ-like helicases than TbRECQ2, indicating that TbRECQ1 is a diverged RecQ-like helicase (or at least more so than TbRECQ2). *T. brucei* and other kinetoplastids are unusual among unicellular organisms by their possession of more than one predicted RecQ-like helicase. TbRECQ2), and its orthologues in *L. major* and *T. cruzi*, grouped relatively closely with the majority of eukaryotic RecQ helicases (such as BLM and WRN), whereas the RECQ1 orthologues clustered more distantly from most eukaryotic RecQ-helicases. Taken together, these data were considered suggestive that TbRECQ1 may have a kinetoplastid-specific function.

### 5.6.2 TbRECQ1 is essential in bloodstream form *T. brucei*

Inability to generate a homozygous *recq1* knockout (*recq1*<sup>-/-</sup>) mutant, despite repeated attempts, and analysis of growth following *RECQ1* RNAi induction strongly indicate that TbRECQ1 is essential in bloodstream form *T. brucei*. *RECQ1* RNAi resulted in cessation of growth by 48 hours post induction. Analysis of endogenously myc-tagged RECQ1 protein levels by western blot showed strong knockdown from 24 hours post induction, with no protein detectable by western blot.

An essential role for RECQ1 is unusual among RecQ proteins. In all single celled organisms thus far examined, the single RecQ protein can be deleted and the cells are viable, as we have seen for *recq2*<sup>-/-</sup> mutants in *T. brucei*. Indeed, even in multicellular organisms, most RecQ proteins are not essential for viability, perhaps due to the ability of other RecQ proteins to partially or completely compensate for the loss of one RecQ protein (Chakraverty & Hickson, 1999; Cobb & Bjergbaek, 2006; Hickson, 2003).

The inability of TbRECQ2 to compensate for the knockdown of TbRECQ1 suggests that TbRECQ1 has one or more functions sufficiently specialised as to make TbRECQ2 unable to compensate for its loss. Though unusual, essentiality of at least one RecQ-like helicase has been demonstrated: RECQ4 is essential for viability in chicken cells (Abe *et al.*, 2011), and RECQL4 is essential for *Drosophila* viability (Capp *et al.*, 2009). Thus, the sequence conservation of kinetoplastid RECQ1 with mammalian RecQ4 may have functional significance

### 5.6.3 RNAi suggests TbRECQ1 has a nuclear function

Knockdown of TbRECQ1 by RNAi resulted in an accumulation of 1N2K cells at 24 hours, indicating cell cycle arrest, followed by accumulation of abnormal cells. 1N >2K cells accumulated in RNAi induced cells at 48 hours, followed by the generation of zoid cells, suggesting a nuclear defect. It is possible that TbRECQ1 plays a role in nuclear DNA replication whereby RNAi induced cells that have not successfully completed nuclear DNA replication proceed through cytokinesis, albeit after some arrest at G2/M, and one of the daughter cells is a zoid. Zoids form as a result of incomplete nuclear DNA replication and consequently, the

nucleus cannot be segregated. FACS analysis of the DNA content of cells and analysis of the proportion of cells undergoing nuclear DNA replication would be useful in testing this hypothesis. Analysis of the incorporation of a thymidine analogue (BrdU or EdU) by fluorescence microscopy is an established technique (Benmerzouga *et al.*, 2013; Wang *et al.*, 2012) for determining whether cells are replicating their DNA.

Further evidence for a nuclear RECQ1 function was seen in that RNAi induction also caused a four to five-fold increase in nuclear  $\gamma$ H2A intensity, as measured by immunofluorescence, indicating an increase in nuclear DNA damage. The nature and source of this damage is unknown, as  $\gamma$ H2A is a general marker of DNA damage and there are multiple sources of DNA damage, such as replication-associated, reactive oxygen species and transcription-associated damage. Cell cycle analysis suggested that RECQ1 may play a role in nuclear DNA replication; RecQ-like helicases in other organisms play roles in all stages of replication and also function in replication-associated events (Croteau *et al.*, 2014). For example, yeast SGS1 (Cobb *et al.*, 2003) interacts with DNA replication proteins and RECQ1 RECQ4/RECQL4 proteins from several organisms play a role in the initiation of replication (Abe *et al.*, 2011; Xu *et al.*, 2009a). RecQ proteins also function in replication fork recovery and stabilization of the replication fork (Sidorova *et al.*, 2013). If TbRECQ1 has a function at any stage of replication or replication-associated repair, DNA lesions resulting from TbRECQ1 knockdown would lead to the observed increase in  $\gamma$ H2A.

Alternatively, TbRECQ1 may not play a role in replication and the source of increased  $\gamma$ H2A expression following TbRECQ1 knockdown is not directly associated with a replication defect. As discussed, RecQ helicases perform a wide range of nucleic acid functions to contribute to genome stability, including DNA repair, recombination and transcription and further investigation is required to dissect the function of TbRECQ1. Many RecQ helicases are involved in the response to induced DNA damage. For example, both BLM and WRN function in the repair of DSBs (Karmakar *et al.*, 2006; Lan *et al.*, 2005). MMS and phleomycin both induce DSBs but knockdown of TbRECQ1 led to no pronounced change in MMS sensitivity, nor was myc epitope-tagged TbRECQ1 observed to form DNA damage foci. Other inducers of DSBs and forms of DNA damage other than DSBs may shed more light on this, but the presented in this chapter do not

suggest that TbRECQ1 plays a profound role in DNA repair. An important consideration, however, in the interpretation of the phleomycin treatment data is that it was not possible to localise TbRECQ1 12myc at all, despite confirming its expression by western blot and its functionality by deletion of the second untagged *TbRECQ1* allele. Further attempts to localise TbRECQ1 would be advantageous, as it would enable co-localisation studies with, for instance, repair associated RAD51 and the DNA damage marker  $\gamma$ H2A, shedding further light on its function. This would determine whether TbRECQ1 is recruited to sites of DNA damage, and therefore whether or not these are replication or repair-associated. It is possible that the 12myc tag used here may be insufficient for IFA localisation and use of an alternative epitope tag such as GFP, HA or FLAG could be attempted. Alternatively, it may be valuable to overexpress RECQ1.

An important further experiment to carry out is transfection of a recoded version of *TbRECQ1* not targeted by the *RECQ1* RNAi fragment. In this recoded cell line, absence of all of the RNAi-associated phenotypes presented in this chapter when RNAi was induced would demonstrate that targeting of TbRECQ1 specifically was the cause of these defects.

Based on the evidence presented in this chapter, we conclude that TbRECQ1 (Tb427.06.3580) is an essential, diverged RecQ-like helicase that likely carries out a nuclear repair or replication function. Further experiments to explore both of these possible roles for TbRECQ1 are necessary.

## **Chapter 6**

### **Discussion**

## 6 Discussion

### 6.1 Introduction

The aim of this project was to analyse whether TbRECQ1, TbRECQ2, TbMUS81 and TbPIF6 are involved in DNA repair in *T. brucei*. Homologues of these genes in other eukaryotes are involved in DNA repair and as some *T. brucei* DNA repair factors are involved in VSG switching, such as RAD51 and BRCA2 (Hartley & McCulloch, 2008; McCulloch & Barry, 1999), it was also aimed to investigate whether these factors are involved in VSG switching. Furthermore, it was aimed to test the prediction that this reaction is initiated by DNA DSBs.

To analyse the functions of the above proteins, it was attempted to generate heterozygote (+/-) and knockout (-/-) mutants for each gene. This was unsuccessful for TbRECQ1 and so the function of this factor was analysed using an inducible RNAi line. The function of the three remaining factors was performed comparing the +/- and -/- cells with wild type *T. brucei*, including in cell lines allowing the induction of DSBs using I-SceI and containing constructs to assay VSG switching. In the latter case, negative selection against a hygromycin thymidine kinase (*HYG-TK*) gene integrated upstream of the active VSG BES acted as a proxy for VSG switching, in which new VSG coat variants are normally selected for via host immune recognition and killing. Though to date the roles of TbRECQ1 and TbMUS81 in VSG switching have not been examined, *Tbmus81* mutants in the VSG switching assay cell line were generated and confirmed by PCR (see Section 4.7.3) and could be used in future work to investigate whether TbMUS81 has a role in VSG switching. In addition, a modified RNAi approach could be considered to test for a role for TbRECQ1, as has been described for other essential factors, such as histone H1, ORC1 (origin recognition complex 1) and TRF (Benmerzouga *et al.*, 2013; Jehi *et al.*, 2014a; Povelones *et al.*, 2012). A summary of the data presented in this thesis and described below is shown in Table 6-1.



| Method                        | Phenotype                            | RECQ1 (RNAi)                           | RECQ2 (-/-)  | MUS81 (-/-)   | PIF6 (-/-)   |
|-------------------------------|--------------------------------------|--|--|---------------|--|
| Cell density                  | <b>Growth</b>                        | Essential                              | Defect   | Defect        | Defect   |
| DAPI                          | <b>Cell cycle</b>                    | 24h: ↑1N2K;<br>48h+: ↑1N>2K<br>+ zoids | None detected  | None detected | None detected  |
| IFA                           | <b>γH2A (+RNAi)</b>                  | ↑                                      | -  | -             | -  |
| Clonal survival               | <b>MMS</b>                           | No change detected                     | ↓survival  | ↓survival     | ↑survival  |
|                               | <b>HU</b>                            | -                                      | ↓survival  | No change     | ↓survival  |
|                               | <b>Phleomycin</b>                    | -                                      | ↓survival(?)   | ↓survival     | No change  |
| I-SceI                        | <b>Survival internal DSB</b>         | -                                      | ↓survival  | ↓survival     | ↓survival  |
|                               | <b>Survival active VSG BES DSB</b>   | -                                      | ↓survival  | ↓survival     | ↑survival  |
|                               | <b>Cell cycle internal DSB</b>       | -                                      | -  | -             | G2/M arrest 24h (as WT)                                  |
|                               | <b>Cell cycle active VSG BES DSB</b> | -                                      | -  | -             | G2/M arrest 24h (as WT) + 'other' cells; intrinsic ↑1N2K |
|                               | <b>γH2A internal DSB</b>             | -                                      | -  | -             | no change from WT  |
|                               | <b>γH2A active VSG BES DSB</b>       | -                                      | -  | -             | Slower resolution of γH2A                                |
| IFA (endogenous 12myc tag)    | <b>Localisation -phleomycin</b>      | Not detected                           | Not detected   | Not detected  | Nuclear  |
|                               | <b>Localisation +phleomycin</b>      | Not detected                           | Nuclear foci; ~80% partial/complete co-localisation with RAD51 | Nuclear foci  | ↑intensity   |
| in vitro; HYGTK reporter gene | <b>VSG switching rate</b>            | -                                      | ↑x2-3  | -             | data inconclusive  |
|                               | <b>VSG switching profile</b>         | -                                      | ↑recombination   | -             | data inconclusive  |

**Table 6-1** Summary of the data presented *recq2*, *mus81* and *pif6* mutants and *recq1* RNAi cells

Arrows indicate increase (↑) or decrease (↓) in value of assayed phenotype; -, experiment not performed; WT, wild type; ?, data unclear; -/-, double knockout mutant.

## 6.2 TbRECQ2 is a DNA repair helicase that suppresses VSG switching

Sequence analyses (see Section 3.2) suggested that TbRECQ2 is a RecQ-like helicase, diverged from other eukaryotic RecQ-like helicases but grouping with

homologues of the *H. sapiens* BLM and its yeast homologue SGS1 in phylogenetic analysis. Phylogenetic analysis also showed that TbRECQ2 is more similar to eukaryotic RecQ-like proteins than the other RecQ-like helicase examined in this thesis, TbRECQ1 (see Chapter 5).

Analysis of *recq2* mutants showed that *recq2*<sup>-/-</sup>, and to some extent *recq2*<sup>+/-</sup> mutants, have a growth defect (see Section 3.4.1). A range of DNA repair defects was also evident in *recq2*<sup>-/-</sup> mutants (see Section 3.4.3), including a striking sensitivity to MMS and HU and some sensitivity to phleomycin when analysed in clonal survival assays, suggesting the involvement of TbRECQ2 in the response and or DNA repair of damage induced by these agents. HU depletes the dNTP pool, leading to replication stalling (Bianchi *et al.*, 1986), which can lead to the formation of DNA breaks. MMS causes widespread damage, stalling replication and resulting in the formation of DSBs (Brookes & Lawley, 1961; Lundin *et al.*, 2005; Wyatt & Pittman, 2006), while phleomycin results in the formation of single and double strand DNA breaks, primarily the latter (Falaschi & Kornberg, 1964; Reiter *et al.*, 1972).

The increased MMS and HU sensitivity observed here in *recq2*<sup>-/-</sup> mutants is consistent with RecQ-like helicases mutants in other organisms. *S. cerevisiae* *sgs1* and human *blm*<sup>-/-</sup> mutants are hypersensitive to MMS (Mullen *et al.*, 2000; Wang *et al.*, 2003), while human *blm*<sup>-/-</sup> and *wrn*<sup>-/-</sup> mutants are involved in replication fork progression and suffer higher levels of DNA damage following HU treatment (Mao *et al.*, 2010; Sidorova *et al.*, 2013). These analyses suggest that TbRECQ2 may localise to replication forks and TbRECQ2 absence, treatment with agents that stall replication (MMS and HU) leads to impaired survival. Analysis of the progression of the replication fork in *recq2*<sup>-/-</sup> mutants following HU or MMS treatment could be used to investigate this further. This could be accomplished using SMARD (single molecular analysis of replicated DNA), for which techniques in trypanosomes have very recently been developed (Calderano *et al.*, 2015).

Endogenously tagged 12myc TbRECQ2 was observed to localise to discrete subnuclear foci following phleomycin treatment, with substantial co-localisation with the HR factor RAD51 (Section 3.9.1). RAD51 foci, not normally observed in the absence of damage, are thought to be repair-related structures (Dobson *et al.*, 2011; Hartley & McCulloch, 2008; Proudfoot & McCulloch, 2005; Trenaman

*et al.*, 2013), suggesting that 12myc RECQ2 is recruited to repair foci in response to phleomycin-induced DNA damage.

It was chosen to look specifically at the role of TbRECQ2 in DSB repair due to the proposed role of DSBs in the current model for the initiation of VSG switching, namely that DSBs in the region of the 70 bp repeats of the active expression site are thought to be the breaks that initiate VSG switching by recombination (Boothroyd *et al.*, 2009; Glover *et al.*, 2013a). To look specifically at the role of TbRECQ2 in the repair of DSBs, the I-SceI meganuclease system was used (see Section 4.2). Analysis of *recq2* mutant survival following the induction of DSBs specifically adjacent to the 70 bp repeats of the active VSG BES and at a chromosomal internal location, showed that *recq2*<sup>-/-</sup> mutants are hypersensitive to DSBs at both of these locations. In both DSB locations, the survival of *recq2*<sup>-/-</sup> cells was ~50% that of wild type, indicating that *recq2*<sup>-/-</sup> mutants are deficient in DSB repair. Importantly, this deficiency also exists when DSBs are induced at the genomic location currently hypothesised as the site for breaks initiating switching.

Analysis of the role of TbRECQ2 in VSG switching, using an *in vitro* VSG switching assay (see Section 4.7), revealed that *recq2*<sup>-/-</sup> mutants exhibit an elevated VSG switching rate. Additionally, there was a dramatic shift in the VSG switching profile of *recq2*<sup>-/-</sup> mutants. The switching profile in wild type cells was dominated by events resulting in the loss of the active VSG BES (~60%) and *in situ* switching events, whereas 90% of surviving *recq2*<sup>-/-</sup> clones underwent VSG gene conversion or telomere crossover events. It seems likely therefore, that TbRECQ2 acts to suppress VSG switching by limiting crossover events. This is markedly similar to the VSG switching phenotype observed in *topo3a* and *rmi1* mutants (Kim & Cross, 2010; Kim & Cross, 2011). Both *topo3a* and *rmi1* mutants display an elevated switching rate, though exceeding that of TbRECQ2 (10-40 and four-fold, respectively), accompanied by an increase in crossover VSG switching events similar to TbRECQ2. Suppressing hyper-recombination is a function of other RecQ-like helicases, such as BLM and WRN (Prince *et al.*, 2001; Yamagata *et al.*, 1998).

These data suggest that TbRECQ2 may be the RecQ helicase component of the RTR (RecQ-Top3-Rmi1) complex, the other two members of which are proposed

by Kim and Cross (2010, 2011) to be Topo3 $\alpha$  and RMI1. A role for TbRECQ2 in the RTR complex seems much more likely than for TbRECQ1, which analysis suggests is more diverged from other eukaryotic RecQ-like helicases than TbRECQ2 and has a nuclear replication rather than DNA repair role (Chapter 5). The RTR complex is a major regulator of crossover in eukaryotes, resolving dHJs to produce non-crossovers (Bussen *et al.*, 2007; Chelysheva *et al.*, 2008; Raynard *et al.*, 2008) and mutants in all three RTR proteins in multiple eukaryotes (human Bloom syndrome cells, *S. cerevisiae* *sgs1*, *top3* and *rmi1* mutants and *S. pombe* *rqh*) display hyper-recombination (Mullen *et al.*, 2005; Myung *et al.*, 2001; Stewart *et al.*, 1997; Traverso *et al.*, 2003; Ui *et al.*, 2001; Wallis *et al.*, 1989). These data suggest that TbRECQ2 may act with RMI1 and TOPO3 $\alpha$  as the RTR complex to suppress the VSG switching rate by limiting crossover events near the VSG in the active VSG BES, suppressing VSG gene conversions and telomere crossover events. *T. brucei* Topo3 $\alpha$  and RMI1 have been demonstrated to interact (Kim & Cross, 2011) and testing for interaction of TbRECQ2 with each of these factors, such as by isolation of an TbRECQ2-Topo3 $\alpha$ -RMI1 complex by co-immunoprecipitation and VSG switching analysis of TbRECQ2 and RMI1 or Topo3 $\alpha$  double mutants would enable this hypothesis to be tested.

The data presented in this thesis on the function of TbRECQ2 in DNA repair and VSG switching also provide new insight into the process of VSG switching by recombination. The DSB repair and VSG switching phenotypes of *recq2*<sup>-/-</sup> mutants described here are not compatible with the hypothesis that the direct formation of DSBs at the 70 bp repeats initiates VSG switching. *recq2*<sup>-/-</sup> mutants are deficient in repair of DSBs induced adjacent to the 70 bp repeats at the active VSG BES, displaying lower survival following DSB induction. That a direct role in the repair of DSBs is a key function of TbRECQ2 is evidenced by the observation that the same reduced survival is seen after I-SceI DSB formation in the interior of chromosome 11. Indeed, the formation of TbRECQ2 foci formation after phleomycin treatment (and known cause of DSBs) suggests that the helicase localises to DSBs, consistent with the I-SceI survival data. If DSBs in the VSG BES directly initiate the VSG switching process, it would be expected that *recq2*<sup>-/-</sup> mutants would display a lower VSG switching rate, as they would be less able to execute the reaction. Instead, the opposite was observed when the VSG switching phenotype of *recq2*<sup>-/-</sup> cells was analysed, since the VSG switching

rate of *recq2*<sup>-/-</sup> cells was two-three-fold higher than wild type cells and occurred by a different route. These data suggest that direct formation of DSBs in the region of the 70 bp repeats is not an adequate explanation for the initiation of VSG switching. Indeed, though the I-SceI system is increasingly being used to study VSG switching (Boothroyd *et al.*, 2009; Glover *et al.*, 2013a; Glover & Horn, 2014), the data presented here suggest that this may not be an accurate mimic of the natural route for initiation of VSG switching.

If DSBs adjacent to the 70 bp repeats are not an accurate model for the initiation of VSG switching, what then is the cause of the lesions detected in the active VSG BES? Several possibilities exist, including replication stalling, transcription-associated instability and the formation of unusual DNA structures (possibly at the 70 bp repeats). Replication stalling within the active VSG BES has been proposed by Kim and Cross (2010 & 2011) as the cause of breaks found in the active VSG BES. However, there is currently no data available on replication dynamics through the active or inactive VSG BESs; for instance, no analysis has been performed to ask if replication forks stall around the VSG, including on the 70 bp repeats. However, this hypothesis is consistent with the role of RecQ-like helicases at the replication fork and in recovery after replication arrest in other organisms (Mao *et al.*, 2010; Sidorova *et al.*, 2013), as well as the sensitivity to the replication stalling agent HU observed in *recq2*<sup>-/-</sup> mutants (see Section 3.4.3.2) and therefore merits further investigation. Increased replication stalling or poorer recovery of the replication fork after replication stalling in the active VSG BES in *recq2*<sup>-/-</sup> mutants, and subsequent formation of lesions, might explain the increased VSG switching rate in *recq2*<sup>-/-</sup> cells. Analysis of replication direction and rate through the active VSG BES using SMARD (single molecule analysis of replicated DNA) could be used to investigate whether replication stalls are seen more often in the active VSG BES than elsewhere in the genome and to compare stalling frequency in wild type cells and *recq2*<sup>-/-</sup> mutants.

### 6.2.1 Summary

The data presented in this thesis demonstrate a clear DNA repair function for *TbRECQ2*. In addition to sensitivity to MMS and HU, mutants are also sensitive to induced DSBs. VSG switching assays also showed that *TbRECQ2* plays a role in

VSG switching, where it appears to suppress the VSG switching rate and inhibit crossovers near the active VSG. The elevated switching rate observed in *recq2*<sup>-/-</sup> mutants, which are also deficient in DSB repair, calls into question the hypothesis that DSBs directly initiate VSG switching by recombination and points to a need to continue to investigate the event and lesion initiating VSG switching by recombination.

### 6.3 *TbMUS81* is a DNA repair factor acting in DSB repair

Amino acid sequence analysis (Section 3.2.2) of TbMUS81 confirmed that it is a MUS81 endonuclease, though somewhat diverged from eukaryotic MUS81 homologues. TbMUS81 contains a predicted ERCC4 domain typical of MUS81 proteins and, in addition, a predicted SAP domain. SAP domains are involved in DNA binding (Aravind & Koonin, 2000) and not found in other eukaryotic MUS81 proteins, suggesting TbMUS81 could possess functions atypical of a MUS81 protein.

Analysis of *mus81* mutants revealed that *mus81*<sup>-/-</sup>, but not *mus81*<sup>+/-</sup> cells, display a minor growth defect (see Section 3.4.1) and analysis of the DNA damage sensitivity of mutants revealed a range of repair defects (see Section 3.4.3). *mus81*<sup>-/-</sup> mutants were more sensitive to MMS and phleomycin, similar to *T. brucei recq2*<sup>-/-</sup> mutants and consistent with MMS sensitivity in *S. cerevisiae mus81Δ* cells (Blanco *et al.*, 2010; Interthal & Heyer, 2000). In contrast with *T. brucei recq2*<sup>-/-</sup> mutants, and to *S. pombe* and *S. cerevisiae mus81Δ* mutants (Boddy *et al.*, 2000; Doe *et al.*, 2002; Doe & Whitby, 2004), *T. brucei mus81*<sup>-/-</sup> cells display at most a minor sensitivity to HU. However, MUS81 has been shown to not be vital for the viability of *S. pombe mus81Δ* cells under acute (as opposed to chronic) HU treatment (Kai *et al.*, 2005) and human *mus81*<sup>-/-</sup> cells are not hypersensitive to HU (Nomura *et al.*, 2007). The HU sensitivity of *mus81* mutant cells may therefore differ by species and by treatment duration, complicating comparison of TbMUS81 with eukaryotic homologues. It would however, be interesting to investigate whether replication fork recovery is impaired in *T. brucei mus81*<sup>-/-</sup> cells, as Mus81 is required for recovery from replication fork collapse in *S. cerevisiae* (which HU-induced dNTP depletion can cause, Petermann *et al.* (2010)).

In addition to a substantial increase in phleomycin sensitivity in *mus81*<sup>-/-</sup> cells, endogenously tagged MUS81 12myc was observed to localise as discrete nuclear foci following phleomycin treatment (see Section 3.9.2), which induces primarily DSBs (Groth *et al.*, 2010; Reiter *et al.*, 1972). These foci are reminiscent of RAD51 DNA repair foci observed in *T. brucei* RAD51 (Dobson *et al.*, 2011; Hartley & McCulloch, 2008; Proudfoot & McCulloch, 2005; Trenaman *et al.*, 2013), suggesting that TbMUS81 is recruited to sites of phleomycin-induced DNA damage. However, the extent of overlap with RAD51 foci was not as striking as that seen for TbRECQ2, for reasons that remain unclear. Nonetheless, together with increased sensitivity in *mus81*<sup>-/-</sup> mutants to phleomycin and MMS, these data suggest that TbMUS81 may be recruited to DSBs and play a role in their repair. Further evidence for a role of TbMUS81 in the repair of DSBs comes from the analysis of cell survival following induced DSBs, using the I-SceI system (see Sections 4.4.2 & 4.5.2). *mus81*<sup>-/-</sup> cells are deficient in DSB repair for DSBs induced at a chromosome-internal location (HR1 cells) and DSBs induced adjacent to the 70 bp repeats of the active VSG BES.

In other eukaryotes, some aspects of MUS81 function and MUS81 substrates are unclear or debated, making comparison with the data on TbMUS81 challenging. MUS81 resolves recombination intermediates arising during replication, replication fork restart and meiosis (Boddy *et al.*, 2001; Hanada *et al.*, 2007; Interthal & Heyer, 2000; Kikuchi *et al.*, 2013; Pepe & West, 2014a), acting on 3' flaps, replication structures and possibly intact HJs (Gaillard *et al.*, 2003; Hollingsworth & Brill, 2004). Failure to repair these endogenously generated recombination intermediates in *mus81*<sup>-/-</sup> mutants could lead to the growth defect observed, and to sensitivity to DNA damaging agents, which lead to an increase in the structures on which TbMUS81 acts. However, further study is required to investigate on which structures TbMUS81 acts. Examination of the repair products following DSB induction in TbMUS81 mutant HR1 and HRES cells would be informative in understanding the function of TbMUS81 in resolving recombination intermediates in *T. brucei*.

The VSG switching rate and profile of *mus81* mutants will be a valuable next experiment to perform. As MUS81 is involved in recovery following replication fork collapse in *S. pombe* (Kai *et al.*, 2005), and given the hypothesised role of replication stalling in the active VSG BES as the initiator of VSG switching,

TbMUS81 may be a key initiator factor, unlike TbRECQ2. As discussed above, MUS81 is also involved in resolving recombination intermediates and therefore *mus81*<sup>-/-</sup> mutants might also be expected to display an altered switching profile as a result of changes in how intermediates are resolved in the absence of MUS81.

### 6.3.1 Summary

The data presented in this thesis suggest that TbMUS81 is involved in the repair of damage induced by phleomycin, MMS and to some extent HU, agents with a wide variety of effects, including replication stalling and DSBs. Further analysis showed TbMUS81 to be important for the repair of DSBs. Review of the literature of MUS81 function in other eukaryotes suggests that TbMUS81 may be acting on recombination intermediates arising from the repair of DNA damage. However, further experiments that examine how DNA lesions are repaired in *T. brucei mus81*<sup>-/-</sup> mutants and whether TbMUS81 is important in replication fork recovery would be informative in understanding how TbMUS81 functions.

## 6.4 *TbPIF6* is an enigmatic DNA repair factor

Of the four factors examined in this thesis, TbPIF6 is the most challenging to understand. Unlike *recq2*<sup>-/-</sup> and *mus81*<sup>-/-</sup> mutants, and cells following *TbRECQ1* RNAi knockdown, *pif6*<sup>-/-</sup> mutants do not display an *in vitro* growth defect (see Section 3.4.1). Also contrasting with *recq2*<sup>-/-</sup> and *mus81*<sup>-/-</sup> cells, which show increased sensitivity to MMS, *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> mutants display an increase in MMS resistance (Section 3.4.3). This increase in MMS resistance in *pif6* mutants does not appear to have been described in any other Pif1 helicase mutant described to date. In fact, other Pif1 helicase mutants, such as in *S. cerevisiae*, are more sensitive to MMS than wild type (Budd *et al.*, 2006), suggesting TbPif6 may be unusual in this respect. In contrast, *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> cell hypersensitivity to HU (Section 3.4.3) is more typical of the phenotype of other Pif1 helicase mutants, such as *S. pombe* Pfh1p mutants (Pinter *et al.*, 2008). It is unusual that the *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> cells displayed similar phenotypes, as this suggests that loss of a single allele is as significant as loss of both alleles, a genetic outcome that has not been observable for other genes in *T. brucei* (no examples were found when conducting a literature search). What this might



mean in terms of the role or abundance of TbPIF6 remains unclear.

TbPIF6 12myc was observed to localise to the nucleus, with an increase in fluorescence intensity following phleomycin damage, consistent with a nuclear repair function. However, it did not form subnuclear foci reminiscent of DNA repair factors such as RAD51, or TbRECQ2 (described above) (see Section 3.9.3). Again then, in what way TbPIF6 contributes to maintenance of the *T. brucei* genome is elusive.

The response to induced DSBs in *pif6* mutants, analysed using the I-SceI system, produced a challenging set of data to interpret, primarily due to inconsistent results upon repetition. Inconsistent results could be due to problems with instability of the I-SceI cell lines, which requires further investigation. However, based upon the initial experiments, which were judged to be more robust, *pif6* mutants are more resistant to DSBs induced adjacent to the 70 bp repeats of the active VSG BES (HRES cells; see Section 4.5.3), which was associated with unexpected retention of VSG221. In contrast, *pif6* mutant were more sensitive to DSBs induced at a chromosomal internal location (HR1 cells; see Section 4.4.3), suggesting that TbPIF6 may function differently depending on the genomic environment. Taking subsequent experiments into consideration, however, would suggest that *pif6* mutants have at least equivalent survival to wild type when DSBs are induced in the active VSG BES. At the very least then, *pif6*<sup>-/-</sup> mutants are not more sensitive to DSBs than wild type cells. DAPI analysis of HRES cells suggested that  $\gamma$ H2A expression, an early marker of DNA damage, dissipates more rapidly following DSB induction in *HRES pif6*<sup>-/-</sup> cells compared to HRES wild type cells (see Section 4.6.3). This could be due to more rapid repair of the DSB or, alternatively, that the cells ignore the damage and prematurely re-enter the cell cycle. Ligation mediated PCR and DNA sequencing could be used to assess whether (and how) the DSB is repaired in *HRES pif6*<sup>-/-</sup> cells. No such rapid decrease in  $\gamma$ H2A compared to wild type was observed in *HR1 pif6*<sup>-/-</sup> cells, again suggesting differences in TbPIF6 function in different genomic environments.

Differences for TbPIF6 DSB repair function at the telomere (HRES) and internal to the chromosome (HR1) could be accounted for if TbPIF6 suppresses telomere addition, similar to activities described for *S. cerevisiae* PIF1. *S. cerevisiae pif1* $\Delta$  mutants have lengthened telomeres and DSBs are repaired by telomere

addition (Li *et al.*, 2014; Makovets & Blackburn, 2009; Schulz & Zakian, 1994). Telomere addition to repair DSBs in HRES cells would be advantageous due to the small amount of sequence lost, but deleterious in HR1 cells due to the large portion of the chromosome that would be lost. Examining the telomere length of *pif6* mutants and investigating whether DSBs are repaired by telomeric addition would determine whether or not TbPIF6 shares this function in *S. cerevisiae* PIF1.

Analysis of the VSG switching phenotype of *pif6* mutants produced an unclear set of data to interpret (see Section 4.7.4.2). The calculated VSG switching rate of *pif6*<sup>-/-</sup> cells was higher than that of wild type and *pif6*<sup>+/-</sup> cells, though it is possible that this was due to a high frequency (50%) of TK mutants in surviving *pif6*<sup>-/-</sup> clones analysed. The switching profile of wild type cells in this experiment was also unusual compared to results obtained in the *RECQ2* VSG switching experiments. These aspects of the data made it difficult to analyse of the VSG switching rates and profiles of *Tbpif6* mutants. Further work is therefore necessary to try to understand the role of *TbPIF6* in VSG switching. To attempt to test for the elevated rate of TK mutants, , further TbPIF6 VSG switching assays should enrich for non-VSG221 expressers following the 48 hour switching period. This could be achieved using a technique such as magnetic activated cell sorting (MACS) utilising an anti-VSG221 antisera, which has been successfully used elsewhere (Boothroyd *et al.*, 2009; Kim & Cross, 2010; Kim & Cross, 2011), or testing for switching rates in immunised mice. Alternatively, sequencing could be used to ask if the ‘switchers’ really are TK mutants, since it remains possible that the elevated rate of ‘switching’ merely reflects altered repair dynamics in PIF6 mutants that affect the *HYG-TK* cassette but do not elicit a change in expressed VSG.

### 6.4.1 Summary

Interpreting the data on the function of *TbPIF6* presented in this thesis is challenging. It is clear that *TbPIF6* provides a nuclear DNA repair function and that this function is quite different from that of TbMUS81 and TbRECQ2. *TbPIF6* appears to be involved in the response to HU treatment, but knockout of *TbPIF6* improves survival following MMS treatment and genomic context may be important in the role of *TbPIF6* in DSB repair. Whether *TbPIF6* plays a role in

VSG switching was unclear, complicated by a high frequency of unswitched cells. Establishing the DSB survival phenotype of *Tbpif6* mutants using several robust experiments would lead to a better understanding of *TbPIF6* DSB repair, while further VSG switching experiments that attempt to remove unswitched cells from analysis would help determine whether *TbPIF6* has a role in VSG switching.

## **6.5 *TbRECQ1* is an essential RecQ helicase in blood stream form cells, with a possible nuclear function**

Analysis of *TbRECQ1* suggests that it is an essential gene in BSF *T. brucei*. Multiple attempts to generate *recq1*<sup>-/-</sup> mutants (see Section 5.3.1) were unsuccessful after multiple attempts and the analysis of a *RECQ1* RNAi cell line showed an arrest in growth at 48 hours post RNAi induction (see Section 5.4.1). *TbRECQ1* RNAi knockdown was shown to be strong, as measured by decrease in *TbRECQ1* 12myc by western blot (Section 5.4.5). An essential role for *TbRECQ1* is unusual among RecQ-like helicases. Most multicellular eukaryotes contain multiple helicases and the loss of one can usually be compensated for by another RecQ helicase (Chakraverty & Hickson, 1999; Cobb & Bjergbaek, 2006; Hickson, 2003), though *RECQL4* is essential in chicken DT40 cells (Abe *et al.*, 2011).

A possible nuclear role for *TbRECQ1* was suggested by analysis of cells following *TbRECQ1* RNAi induction by DAPI staining for cell cycle analysis, and immunolocalisation (see Sections 5.4.2 & 5.4.4) for investigation of levels of the DNA damage marker  $\gamma$ H2A. These analyses revealed an increase in the 1N2K cell population at 24 hours post RNAi induction, the appearance of 1N >2K cells at 48 hours, followed by the generation of zoid cells. *TbRECQ1* RNAi induction resulted in a five-fold increase in  $\gamma$ H2A intensity, indicating higher levels of DNA damage following *TbRECQ1* knockdown. *TbRECQ1* may therefore play a role in nuclear DNA replication, indicated by the accumulation of 1N2K cells followed by cells that have not completed nuclear DNA replication (1N >2K and zoid cells). Defects in nuclear replication, such as stalled replication forks, single strand breaks (SSBs) and DSBs may result in the formation of  $\gamma$ H2A, signaling DNA damage. Unfortunately, it was not possible to visualise *TbRECQ1* 12myc localisation in order to test whether *RECQ1* is present in the nucleus (see Section 5.5.4). It is also possible that *TbRECQ1* is not involved in nuclear DNA replication and that there is a different cause for the increase in  $\gamma$ H2A

expression following TbRECQ1 RNAi knockdown. Analysis of the growth of cells following TbRECQ1 RNAi knockdown in the presence of MMS, revealed no detectable phenotype, suggesting that TbRECQ1 does not play a role in the response to MMS-induced DNA damage. A role in the response to other DNA damaging agents cannot be ruled out however.

The essentiality and nuclear defect of TbRECQ1 is reminiscent of several other *T. brucei* DNA repair factors that are essential but not core to replication. PrimPol (primase-polymerase) is a human protein with both primase, to synthesise RNA primers to prime DNA synthesis, and DNA polymerase activities and acts in translesion DNA synthesis (García-Gómez *et al.*, 2013). Translesion DNA synthesis permits replication through replication fork-blocking DNA lesions. *T. brucei* encodes two PrimPol-like (PPL) proteins, PPL1 and PPL2, of which PPL2 is essential in BSF *T. brucei* (Rudd *et al.*, 2013). RNAi knockdown of PPL2 results in G2 cell cycle arrest following DNA synthesis, accumulation of  $\gamma$ H2A and Rad51 foci (Rudd *et al.*, 2013). From these data, Rudd *et al.* (2013) concluded that PPL2 has an essential role in the post-replication tolerance of endogenous DNA damage. Thus, PPL2 while being an essential factor involved in DNA repair, is not a core replication protein. Similarly, the *T. brucei* nucleotide excision repair (NER) genes XPC (xeroderma pigmentosum group C) and DDB (DNA-damage binding) are essential by RNAi knockdown, but are not core replication proteins (Machado *et al.*, 2014). These examples illustrate that it is possible that TbRECQ1 could act in a replication-related but unknown genome maintenance function.

Several lines of investigation suggest that TbRECQ1 could play a kinetoplastid-specific role. Firstly, primary amino acid sequence analysis (see Section 5.2) showed that *T. brucei*, *L. major* and *T. cruzi* each contain two predicted RecQ-like helicases, whereas most unicellular eukaryotes possess a single RecQ-like helicase. In these organisms, one RecQ-like helicase (including TbRECQ2) clusters with eukaryotic BLM/SGS1 orthologues and the other groups more distantly, with TbRECQ1. Primary amino acid sequence analysis also showed that TbRECQ1 appears to be substantially diverged from other eukaryotic RecQ helicases, and more so than TbRECQ2, possessing fewer of the domains typically found in RecQ-like helicases. Whether TbRECQ1 carries out an essential function specific to kinetoplastids requires further study however. Important to this is

that the RecQ-like helicases predicted in other kinetoplastids have not been studied. Establishing whether the phenotype of the TbRECQ1 homologues in other kinetoplastids are similar to that in *T. brucei* would be useful in determining whether TbRECQ1 carries out a kinetoplastid-specific function.

Though the data presented in this thesis suggest an essential nuclear role for TbRECQ1, possibly in DNA replication, further work is required to establish the role of TbRECQ1 in *T. brucei*. To draw a more detailed picture of the cell cycle profile of TbRECQ1 RNAi induced cells, FACS analysis could be used and measurement of the incorporation of a thymidine analogue (e.g. BrdU or EdU) could be measured by immunofluorescence to determine whether cells continue to replicate their nuclear DNA following TbRECQ1 RNAi induction. Use of another epitope tag to immunolocalise TbRECQ1, such as HA, may be more successful than the 12myc tag used in Chapter 5 (see Section 5.5.4) and localisation of TbRECQ1 to the nucleus and testing whether localisation changes in response to replication stalling (e.g. by HU treatment) or by cell cycle stage would be informative in establishing a nuclear replication role for TbRECQ1.

One of the aims at the outset of this thesis was to investigate whether the four factors examined play a role in VSG switching. This was not examined for TbRECQ1, but if TbRECQ1 is involved in replication it does not preclude its involvement in VSG switching. RNAi knockdown of the replication factor ORC1/CDC6 in BSF *T. brucei* results in an increase in the VSG switching rate and de-repression of VSGs in inactive VSG BESs, due to VSG BES instability (Benmerzouga *et al.*, 2013). Another possible mechanism by which a TbRECQ1 could contribute to VSG switching, through a replicative function, is via replication stalling. Replication stalling within the active VSG BES has been hypothesised as the cause of breaks found in the active VSG BES that are thought to initiate VSG switching (Kim & Cross, 2010; Kim & Cross, 2011). Therefore, if TbRECQ1 RNAi knockdown were to cause increased replication stalling, this might affect the VSG switching rate. However, no direct evidence exists for replication stalling in the active VSG BES and as discussed above, more work is needed to understand what, if any, role TbRECQ1 plays in DNA replication.

It is also possible that TbRECQ1 may not play a role in replication and that increased  $\gamma$ H2A expression observed following *TbRECQ1* RNAi is not related to a

replication defect. RecQ helicases perform a range of functions in the genome in addition to replication, including recombination, transcription and DNA repair. Growth curve analysis of MMS-treated *TbRECQ1* RNAi induced cells did not result in a detectable phenotype but this does not exclude the possibility that cells may be sensitive to other forms of DNA damage following *TbRECQ1* RNAi knockdown. Some DNA repair proteins are recruited to sites of DNA damage. As such, successful localisation of TbRECQ1 would be advantageous in this respect, as it would allow the investigation of whether it too is recruited to sites of DNA damage induced by various agents such as MMS and phleomycin.

### 6.5.1 Summary

The data presented here on the function of TbRECQ1 show a diverged RecQ-like helicase, unusual due to its apparent essentiality in an organism with more than one RecQ helicase and with an interesting nuclear phenotype. More work is required to dissect its role in *T. brucei*: as a first step, the hypothesis that it is involved in nuclear DNA replication can be tested by measuring replication in using thymidine analogue incorporation and more detailed analysis of the cell cycle stage distribution of TbRECQ1 RNAi induced cells.

## 6.6 Conclusion

This work has explored the functions in DNA repair of four previously uncharacterised proteins in bloodstream form *T. brucei*. Three of these, TbRECQ2, TbMUS81 and TbPIF6 are DNA repair factor whereas the fourth, TbRECQ1, is essential appears to be involved in nuclear DNA replication. Though all involved in DNA repair, probing of their functions using a range of DNA damaging agents and induced DSBs reveal that TbRECQ2, TbMUS81 and TbPIF6 all play different roles; while *Tbrecq2* and *Tbmus81* mutants are sensitive to several DNA damaging agents including MMS and are sensitive to induced DSBs, *Tbpif6* mutants are resistant to MMS-induced DNA damage and display no increased sensitivity to DSB repair. It was also found that TbRECQ2 plays a role in VSG switching, suppressing VSG switching events in the active VSG BES, a finding that cannot easily be reconciled with the hypothesis that DSBs in the active VSG BES initiate VSG switching. Further studies to investigate the role in DSB repair and VSG switching of TbRECQ2, and utilising the cell lines developed in the course of

this work to investigate VSG switching in *Tbmus81* mutants, would be valuable in furthering our understanding of this process. TbPIF6 provides an intriguing avenue for further study, as an MMS-resistant phenotype has not been observed previously in PIF1-like helicase mutant. Unfortunately, experiments here were unable to determine whether TbPIF6 plays a role in VSG switching but refinements to experiment as discussed above would be expected to yield improved results on this. As demonstrated by data presented here for the role of TbRECQ2 in DNA repair and VSG switching, the investigation of VSG switching by recombination in *T. brucei* would benefit from both examining DSB repair in the active VSG BES in addition to a non-directed approach, in order to continue to assess the hypothesis that DNA breaks in the active VSG BES initiate the process.

# Appendices



## 7 Appendices

### 7.1 Oligonucleotide sequences

| Primer # | Sequence (5'-3')                                  | Restriction enzyme sites |
|----------|---|--------------------------|
| 13       | GTGGCGCGCCAGCCAAATATTCGTTCTGTG                    | Ascl                     |
| 14       | CAGTCTAGACGTTTCTTCTTTTCGTCCA                      | XbaI                     |
| 19       | GTGGCGCGCCGTGATCAAGGTGTCGTTGT                     | Ascl                     |
| 20       | CAGTCTAGAATCATCGTCCACCATTAGC                      | XbaI                     |
| 22       | CAGACTAGTTCTGTCCACAGAATTCAT                       | SpeI                     |
| 23       | CAGGGTACCAGGACAAAACACTAAAAAATA                    | KpnI                     |
| 24       | CAGGGTACCGACAAAGATTTAAGTTGCGTCT                   | KpnI                     |
| 25       | CAGGGATCCTCGCCGCGGTAATAGTTG                       | BamHI                    |
| 26       | GATCTTCAAGCTTGCGGCCGTACAGAGGAAGAGAAGGAGGAAG       | HindIII                  |
| 27       | GATCTTCTCTAGAGGGAGAAACGACCACACCAAATC              | XbaI                     |
| 28       | GATCTTCGAGCTCACTTACACCACACAACCTTGGACG             | SacI                     |
| 29       | GATCTTCATCGATGCGGCCGCGTTTCAGGTAAATGGGCTGTTT       | ClaI                     |
| 30       | GATCTTCAAGCTTGCGGCCGCTGTGTAAATCCGTTCTTTCTTC       | HindIII                  |
| 31       | GATCTTCTCTAGATACAACGACACAATACCAACCAC              | XbaI                     |
| 32       | GATCTTCGAGCTCACAGACAATCTCCATCAGCAACC              | SacI                     |
| 33       | GATCTTCATCGATGCGGCCGCATAAGACATCCACCAGAACCTGC      | ClaI                     |
| 34       | GATCTTCAAGCTTGCGGCCGCGTGAGAAAAGGAAGTGGG           | HindIII                  |
| 35       | GATCTTCTCTAGATAATTGGCAGTGGATATCG                  | XbaI                     |
| 36       | GATCTTCGAGCTCGACGAAGCCACAAAAAGAA                  | SacI                     |
| 37       | GATCTTCATCGATGCGGCCGCCGAAGACAGCAACACCAA           | ClaI                     |
| 38       | GATCTTCAAGCTTGCGGCCGCACACATTGAACAAATACGCTG        | HindIII                  |
| 39       | CTTCTTCTCTAGATGTTTCGTAGTCCAAATGGGTATC             | XbaI                     |
| 40       | GATCTTCGAGCTCAAAGAGAATGTTGGTGGAAGAGC              | SacI                     |
| 41       | GATCTTCATCGATGCGGCCGCAATCCGCTTACCTATCAAGTCC       | ClaI                     |
| 44       | CCCGCCAGAGTATCGTAA                                |                          |
| 45       | TCGCCCAGAATAAGCTCA                                |                          |
| 46       | CCAGCAGAGGGAAAGGAA                                |                          |
| 47       | CCTCCACAACCCACACAA                                |                          |
| 48       | AAACGGGAATGACACTGA                                |                          |
| 49       | CAAAGGAAACACGACAC                                 |                          |
| 50       | CTCTCCTTTTACCACCACTTT                             |                          |
| 51       | GTTGTGTTTTCTTTCCCTG                               |                          |
| 52       | CCCTTGTTAAGCACACCT                                |                          |
| 53       | TACACATTTACCCACAC                                 |                          |
| 67       | GGGGACAAGTTTGTACAAAAAGCAGGCTTCGCTTCGTTGAGGTTTCTT  | AttB site                |
| 68       | GGGGACCACTTTGTACAAGAAAGCTGGGTAGTGGGGAGAACGCACATAC | AttB site                |
| 75       | CCTATCAGCGAGCAAACTC                               |                          |
| 76       | TTCCTCGCTTTTCCATTTG                               |                          |
| 77       | TTTGTGATAACTGCGCAAGC                              |                          |
| 78       | ACCTTGAGTGACGTGAACC                               |                          |
| 79       | CTGTTGTGTGGGTGTGGAG                               |                          |
| 80       | AACGTTTCTGGGCATTGTTC                              |                          |
| 81       | GGTGGGTGTACGATCCATTC                              |                          |

**Table 7-1** List of oligonucleotides used in this thesis

|        |  |                              |
|--------|--|------------------------------|
| 82     | TCGCCAAGGAGAATAACCTG                   |                              |
| 86     | GATCAAGCTTATGGTGCCTCCTCCAAG            |                              |
| 87     | GATCGCTAGATCAGGCACCGGGCTTGC            |                              |
| 90     | GGAGACGAAGAGCCGGTT                     |                              |
| 91     | CCGCTGCAGAAAGCTGGC                     |                              |
| 92     | AATGGAAGAGCAAAGCTGATAGGTTGG            |                              |
| 93     | GGCGGCCACTCCATTGTCTG                   |                              |
| 94     | AATGGAATGGGGAGAGTG                     |                              |
| 95     | GGAAGAACTTGGGCTAGGA                    |                              |
| 108    | CGCTATCTCGAGATGAAAAAGCCTGAACTCACC      | XhoI                         |
| 109    | CGCTATATCGATCTATTCTTTGCCCTCGG          | Clal                         |
| 116    | CTAGCAGCTTCTAGAGCGGGCCGCTCTAGAGGATCCT  | HindIII-NheI-NotI-XbaI-BamHI |
| 117    | CCGGAAGATCTCTAGAGCGGGCCGCGCTAGCAAGCTTG | BamHI-XbaI-NotI-NheI-HindIII |
| 120    | GTGACCACCCTGACCTAC                     |                              |
| 121    | GCAAAACTGTGATGACCCGC                   |                              |
| 126    | CGCCATGCTCTAGATCACTGCGAATCAACTGTGA     | XbaI                         |
| 128    | CGCCAGTGCTCTAGATTAGCAGTAAGTCTCGTCAATA  | XbaI                         |
| 130    | CGCCAGTGCTCTAGATCAATCATCGTCCACCAT      | XbaI                         |
| 131    | TTTACGGGCTACTTGCCAAT                   |                              |
| 133    | CCTCATTTTCTGGATTTTGCTC                 |                              |
| 151    | CGCTATGCTAGCATGGACAAAGATTTAAGTTGCGTC   | NheI                         |
| 152    | CGCTATGCTAGCATGCCGAACAGCCATAACGA       | NheI                         |
| 153    | CGCTATGCTAGCATGGTAGTACTGCGACCA         | NheI                         |
| 154    | GGGTGGATTCTTCTTGAGAC                   |                              |
| 155    | GCGTGCAATCCATCTTGTTT                   |                              |
| 156    | GCAAGTATATACGCTGAAATAAATCAC            |                              |
| 157    | TGTTTGGCTGTTTCGCTACTGTGAC              |                              |
| 158    | CTGGATCCAGCGCCGTTCCACGCGAGA            |                              |
| 159    | GACTCGAGCTATCCCCAATCCGTGCCGTCCCG       |                              |
| 160    | CTTCTTCAAGTCCGCCAT                     |                              |
| 161    | GCTCAGGTAGTGTTGTC                      |                              |
| OL4038 | TAAGGTCTCGTTGCTGCC                     |                              |
| OL4161 | TAATGCCAACTTTGTACAAA                   |                              |
| OL4212 | TAATGCCAACTTTGTACAAG                   |                              |

Table 7-1 continued

## 7.2 Protein accession numbers

| Figure | Sequence name | Organism                         | Accession number | Database             |
|--------|---------------|----------------------------------|------------------|----------------------|
| 3-2    | TbRECQ2       | <i>Trypanosoma brucei</i>        | Tb427.08.6690    | TriTrypDB            |
|        | HsBLM         | <i>Homo sapiens</i>              | NP_000048.1      | NCBI                 |
|        | HsRECQ1       | <i>Homo sapiens</i>              | NP_116559        | NCBI                 |
|        | SpRQH1        | <i>Schizosaccharomyces pombe</i> | CAA91177         | GenBank              |
|        | ScSGS1        | <i>Sachharomyces cerevisiae</i>  | NP_013915.1      | NCBI                 |
|        | HsRECQ5       | <i>Homo sapiens</i>              | O94762           | UniProtKB/Swiss-Prot |
|        | HsWRN         | <i>Homo sapiens</i>              | NP_000544        | NCBI                 |
|        | TbRECQ1       | <i>Trypanosoma brucei</i>        | Tb427.06.3580    | TriTrypDB            |
|        | HsRECQ4       | <i>Homo sapiens</i>              | BAA86899         | GenBank              |
| 3-3    | TcRECQ(2)     | <i>Trypanosoma cruzi</i>         | TcCLB.506941.90  | TriTrypDB            |
|        | TbRECQ1       | <i>Trypanosoma brucei</i>        | Tb427.06.3580    | TriTrypDB            |
|        | LmRECQ(2)     | <i>Leishmania major</i>          | LmjF.30.2290     | TriTrypDB            |
|        | AtRQL2        | <i>Arabidopsis thaliana</i>      | NP_174421        | NCBI                 |
|        | HsRECQ1       | <i>Homo sapiens</i>              | NP_116559        | NCBI                 |
|        | AtRQL3        | <i>Arabidopsis thaliana</i>      | AEE86556         | GenBank              |
|        | AtRQL1        | <i>Arabidopsis thaliana</i>      | CAC14163         | GenBank              |
|        | HsRECQ5       | <i>Homo sapiens</i>              | O94762           | UniProtKB/Swiss-Prot |
|        | AtRQL4B       | <i>Arabidopsis thaliana</i>      | Q9FT70           | UniProtKB/Swiss-Prot |
|        | AtRQL4A       | <i>Arabidopsis thaliana</i>      | Q8L840           | UniProtKB/Swiss-Prot |
|        | HsBLM         | <i>Homo sapiens</i>              | NP_000048.1      | NCBI                 |
|        | ScSGS1        | <i>Sachharomyces cerevisiae</i>  | NP_013915        | NCBI                 |
|        | TcRECQ(1)     | <i>Trypanosoma cruzi</i>         | TcCLB.507433.9   | TriTrypDB            |
|        | TbRECQ2       | <i>Trypanosoma brucei</i>        | Tb427.08.6690    | TriTrypDB            |
|        | LmRECQ        | <i>Leishmania major</i>          | LmjF.24.1530     | TriTrypDB            |
|        | HsWRN         | <i>Homo sapiens</i>              | NP_000544        | NCBI                 |
|        | AtRQSIM       | <i>Arabidopsis thaliana</i>      | Q9FT69           | UniProtKB/Swiss-Prot |
|        | HsRECQ4       | <i>Homo sapiens</i>              | BAA86899         | GenBank              |
|        | AtRQL5        | <i>Arabidopsis thaliana</i>      | Q0WVW7           | UniProtKB/Swiss-Prot |
| 3-5    | TbMUS81       | <i>Trypanosoma brucei</i>        | Tb427.08.6740    | TriTrypDB            |
|        | HsMUS81       | <i>Homo sapiens</i>              | NP_079404.3      | NCBI                 |
|        | SpMUS81       | <i>Schizosaccharomyces pombe</i> | XP_001713161.1   | NCBI                 |
|        | ScMUS81       | <i>Sachharomyces cerevisiae</i>  | Q04149           | UniProtKB/Swiss-Prot |
|        | HsXPF         | <i>Homo sapiens</i>              | Q92889.3         | UniProtKB/Swiss-Prot |
|        | ScRAD1        | <i>Sachharomyces cerevisiae</i>  | P06777.1         | UniProtKB/Swiss-Prot |
|        | TbXPF         | <i>Trypanosoma brucei</i>        | Tb927.5.3670     | TriTrypDB            |
|        | HsEME1        | <i>Homo sapiens</i>              | AAH16470.1       | GenBank              |
|        | ScMUS81       | <i>Sachharomyces cerevisiae</i>  | Q04149.1         | UniProtKB/Swiss-Prot |
|        | SpMUS81       | <i>Schizosaccharomyces pombe</i> | P87231.2         | UniProtKB/Swiss-Prot |
|        | HsMUS81       | <i>Homo sapiens</i>              | NP_079404.3      | NCBI                 |
|        | TcMUS81       | <i>Trypanosoma cruzi</i>         | TcCLB.510065.20  | TriTrypDB            |
|        | TbMUS81       | <i>Trypanosoma brucei</i>        | Tb427.08.6740    | TriTrypDB            |

**Table 7-2** Accession numbers for sequences used in sequence analysis

|     |           |                                  |                  |                      |
|-----|-----------|----------------------------------|------------------|----------------------|
|     | LmMUS81   | <i>Leishmania major</i>          | LmjF.24.1520     | TriTrypDB            |
|     | SpEME1    | <i>Schizosaccharomyces pombe</i> | Q9C103.2         | UniProtKB/Swiss-Prot |
|     | ScMMS4    | <i>Sachharomyces cerevisiae</i>  | P38257.2         | UniProtKB/Swiss-Prot |
| 3-8 | TbPIF6    | <i>Trypanosoma brucei</i>        | Tb427.10.910     | TriTrypDB            |
|     | ScPIF1    | <i>Sachharomyces cerevisiae</i>  | P07271.2         | UniProtKB/Swiss-Prot |
|     | SpPIF1    | <i>Schizosaccharomyces pombe</i> | NP_596488        | NCBI                 |
|     | ScRRM3    | <i>Sachharomyces cerevisiae</i>  | EWB18068         | GenBank              |
|     | HsPIF1    | <i>Homo sapiens</i>              | NP_001273425.1   | NCBI                 |
| 3-9 | TbPIF3    | <i>Trypanosoma brucei</i>        | Tb427tmp.01.3660 | TriTrypDB            |
|     | LmPIF3    | <i>Leishmania major</i>          | LmjF.28.2530     | TriTrypDB            |
|     | TbPIF4    | <i>Trypanosoma brucei</i>        | Tb427tmp.01.6420 | TriTrypDB            |
|     | TcPIF4    | <i>Trypanosoma cruzi</i>         | TcCLB.503677.20  | TriTrypDB            |
|     | LmPIF2    | <i>Leishmania major</i>          | LmjF.11.0320     | TriTrypDB            |
|     | TbPIF1    | <i>Trypanosoma brucei</i>        | Tb427tmp.02.4730 | TriTrypDB            |
|     | LmPIF1    | <i>Leishmania major</i>          | LmjF.11.0330     | TriTrypDB            |
|     | LmPIF4    | <i>Leishmania major</i>          | LmjF.32.1590     | TriTrypDB            |
|     | ScPIF1    | <i>Sachharomyces cerevisiae</i>  | P07271.2         | UniProtKB/Swiss-Prot |
|     | SpPIF1    | <i>Schizosaccharomyces pombe</i> | NP_596488        | NCBI                 |
|     | TbPIF8    | <i>Trypanosoma brucei</i>        | Tb427.07.1000    | TriTrypDB            |
|     | HsPIF1    | <i>Homo sapiens</i>              | NP_001273425.1   | NCBI                 |
|     | TbPIF5    | <i>Trypanosoma brucei</i>        | Tb427.08.3560    | TriTrypDB            |
|     | TbPIF7    | <i>Trypanosoma brucei</i>        | Tb427.08.700     | TriTrypDB            |
|     | TcPIF7    | <i>Trypanosoma cruzi</i>         | TcCLB.508567.80  | TriTrypDB            |
|     | LmPIF7    | <i>Leishmania major</i>          | LmjF.07.1020     | TriTrypDB            |
|     | TbPIF6    | <i>Trypanosoma brucei</i>        | Tb427.10.910     | TriTrypDB            |
|     | LmPIF6    | <i>Leishmania major</i>          | LmjF.21.1190     | TriTrypDB            |
|     | TcPIF6    | <i>Trypanosoma cruzi</i>         | TcCLB.511133.4   | TriTrypDB            |
|     | TbPIF8    | <i>Trypanosoma brucei</i>        | Tb427.07.1000    | TriTrypDB            |
|     | TbPIF2    | <i>Trypanosoma brucei</i>        | Tb427tmp.02.4740 | TriTrypDB            |
|     | LmPIF8    | <i>Leishmania major</i>          | LmjF.26.0930     | TriTrypDB            |
| 5-2 | MmRECQ4   | <i>Mus musculus</i>              | Q75NR7           | UniProtKB/Swiss-Prot |
|     | HsRECQ4   | <i>Homo sapiens</i>              | BAA86899         | GenBank              |
|     | DmRECQ4   | <i>Drosophila melanogaster</i>   | AAF42939         | GenBank              |
|     | AtRQL4B   | <i>Arabidopsis thaliana</i>      | Q9FT70           | UniProtKB/Swiss-Prot |
|     | AtRQL4A   | <i>Arabidopsis thaliana</i>      | Q8L840           | UniProtKB/Swiss-Prot |
|     | DdBLM     | <i>Dictyostelium discoideum</i>  | XP_629849.1      | NCBI                 |
|     | MmBLM     | <i>Mus musculus</i>              | NP_001035992     | NCBI                 |
|     | HsBLM     | <i>Homo sapiens</i>              | NP_000048.1      | NCBI                 |
|     | DmBLM     | <i>Drosophila melanogaster</i>   | NP_524319.2      | NCBI                 |
|     | CeBLM     | <i>Caenorhabditis elegans</i>    | O18017           | UniProtKB/Swiss-Prot |
|     | ScSGS1    | <i>Sachharomyces cerevisiae</i>  | NP_013915        | NCBI                 |
|     | TcRECQ(1) | <i>Trypanosoma cruzi</i>         | TcCLB.507433.9   | TriTrypDB            |
|     | TbRECQ2   | <i>Trypanosoma brucei</i>        | Tb427.08.6690    | TriTrypDB            |

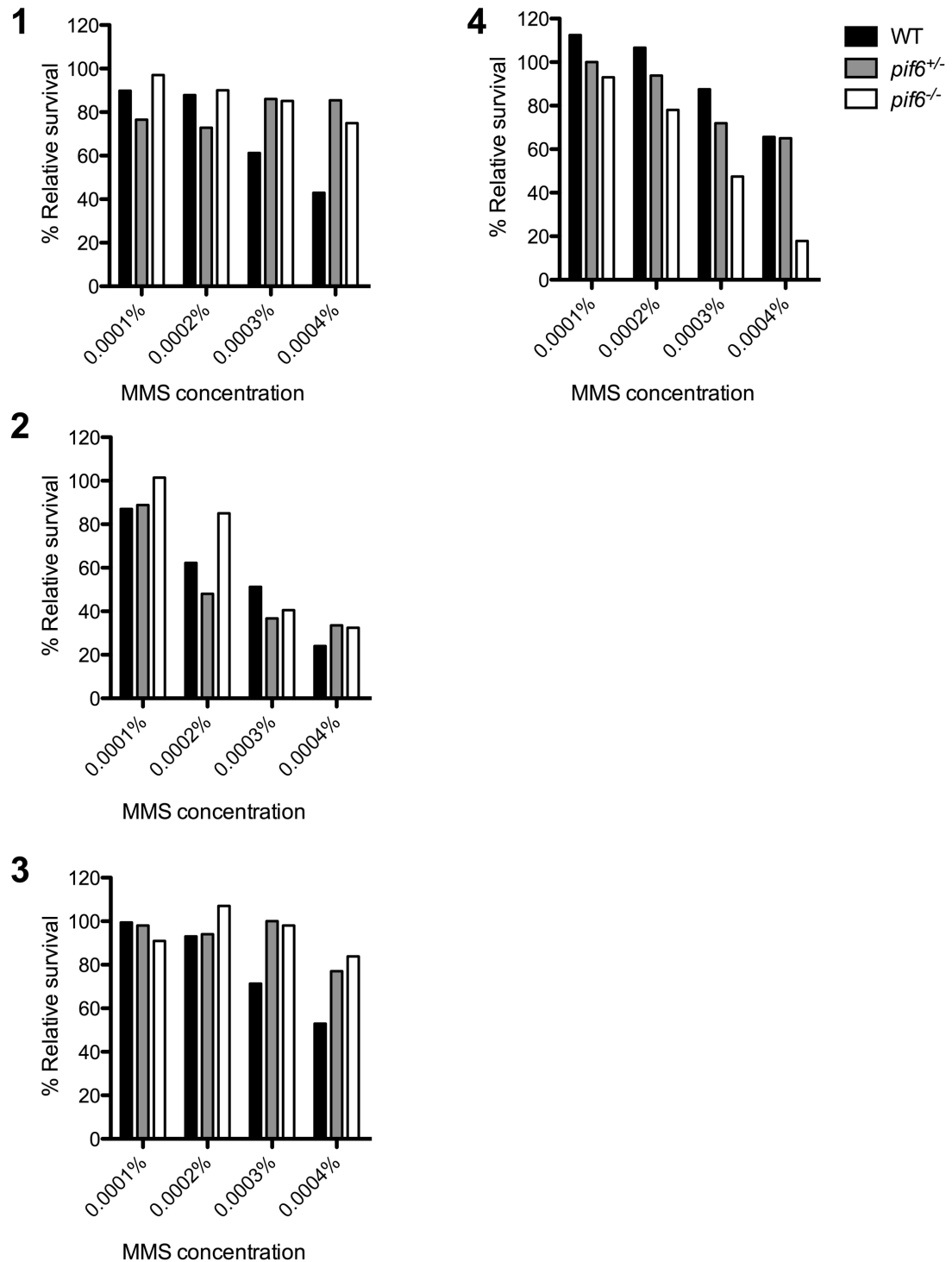
Table 7-2 continued

|           |                                     |                |                      |
|-----------|-------------------------------------|----------------|----------------------|
| LmRECQ(1) | <i>Leishmania major</i>             | LmjF.24.1530   | TriTrypDB            |
| TtRECQ    | <i>Tetrahymena thermophila</i>      | XP_001015163.2 | NCBI                 |
| AeRECQ    | <i>Actinobacillus equuli</i>        | WP_039197383.1 | NCBI                 |
| HspRECQ   | <i>Haemophilus sputorum</i>         | WP_007523284.1 | NCBI                 |
| MhRECQ    | <i>Mannheimia haemolytica</i>       | KIX31114.1     | GenBank              |
| EcRECQ    | <i>Escherichia coli</i>             | AAA24517.1     | GenBank              |
| AaRECQ    | <i>Anditalea andensis</i>           | WP_035071022.1 | NCBI                 |
| BtRECQ    | <i>Bifidobacterium thermophilum</i> | WP_033492071.1 | NCBI                 |
| AtRQL2    | <i>Arabidopsis thaliana</i>         | NP_174421      | NCBI                 |
| HsRECQ1   | <i>Homo sapiens</i>                 | NP_116559      | NCBI                 |
| AtRQL3    | <i>Arabidopsis thaliana</i>         | AEE86556       | GenBank              |
| AtRQL1    | <i>Arabidopsis thaliana</i>         | CAC14163       | GenBank              |
| GiSGS1    | <i>Giardia intestinalis</i>         | XP_001708411.1 | NCBI                 |
| CeQ19046  | <i>Caenorhabditis elegans</i>       | Q19046         | UniProtKB/Swiss-Prot |
| AtRQL5    | <i>Arabidopsis thaliana</i>         | Q0WVW7         | UniProtKB/Swiss-Prot |
| HsRECQ5   | <i>Homo sapiens</i>                 | O94762         | UniProtKB/Swiss-Prot |
| HsWRN     | <i>Homo sapiens</i>                 | NP_000544      | NCBI                 |
| CeWRN     | <i>Caenorhabditis elegans</i>       | NP_495324      | NCBI                 |
| TbRECQ1   | <i>Trypanosoma brucei</i>           | Tb427.06.3580  | TriTrypDB            |
| AtRQSIM   | <i>Arabidopsis thaliana</i>         | Q9FT69         | UniProtKB/Swiss-Prot |
| LmRECQ(2) | <i>Leishmania major</i>             | LmjF.30.2290   | TriTrypDB            |
| VbRECQ    | <i>Vibrio brazilienses</i>          | XP_001566851.1 | NCBI                 |
| PfRECQ    | <i>Plasmodium falciparum</i>        | PF3D7_1429900  | EupathDB             |

Table 7-2 continued

### 7.3 Individual *pif6* mutant MMS clonal survival assays

The three individual *pif6* mutant clonal survival assays following MMS damage, used to calculate the mean relative survival of *pif6* mutants shown in Figure 3-17 (Section 3.4.3.2), are shown in Figure 7-1. Also shown is an additional assay (Fig. 7-1(graph 4)) where *pif6*<sup>-/-</sup> survival was substantially lower than initial assays. This experiment was not included in data shown in Figure 3-7 as it was discounted as an anomaly due to the dissimilar phenotype of *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> mutants. *pif6*<sup>+/-</sup> and *pif6*<sup>-/-</sup> exhibited similar phenotypes in all previous MMS assays conducted. As described in Section 3.4.3.2, *pif6* mutant survival following MMS treatment appeared to decrease in experiments conducted after cells had been in culture for a longer period of time.



**Figure 7-1 Individual *pif6* mutant MMS clonal survival assays**

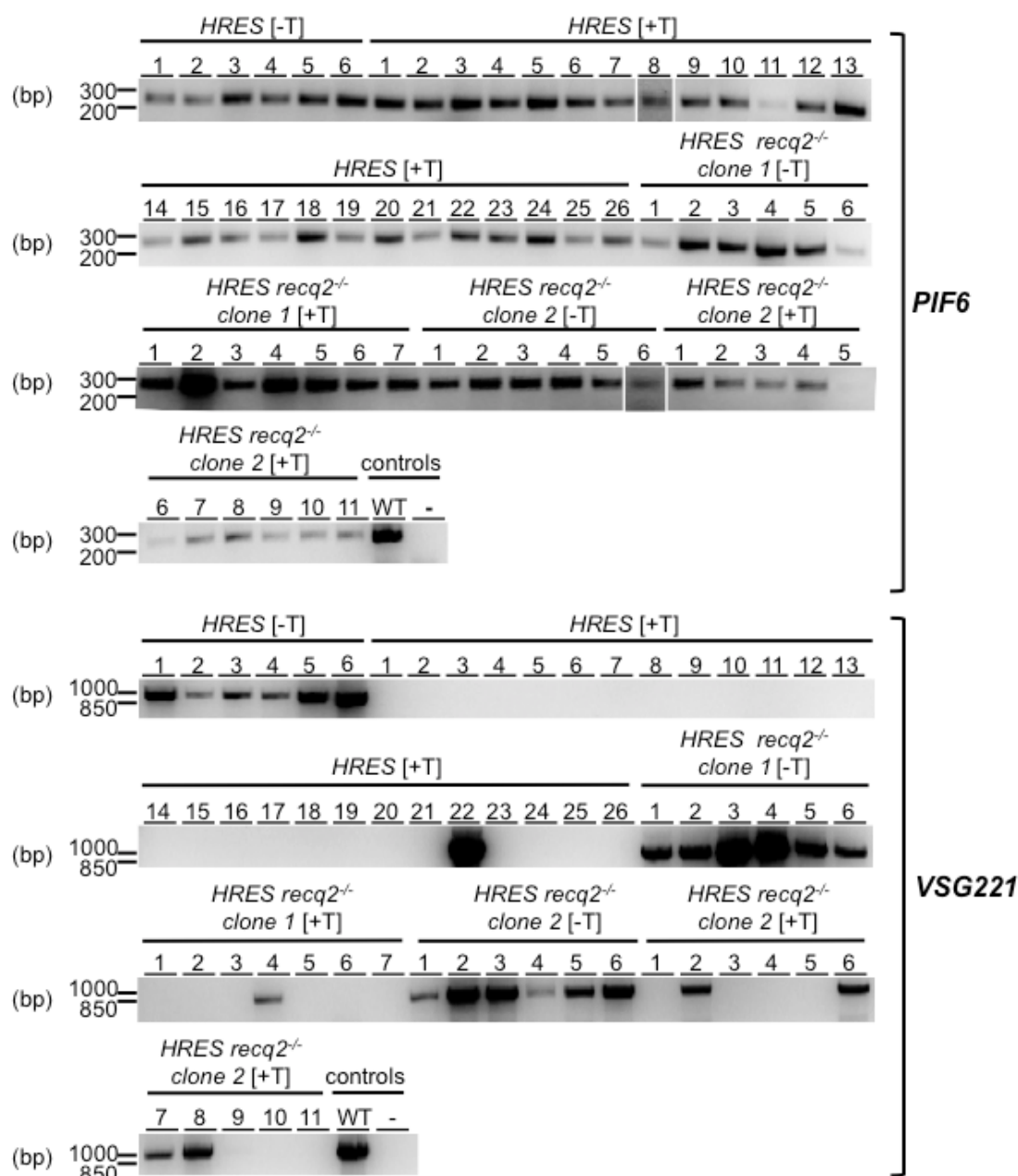
Individual assays of wild type cells and *pif6* mutant survival following MMS damage are shown in chronological order (i.e. 1, first assay; 4, final assay). Assays 1-3 were used to calculate the mean relative survival of wild type and *pif6* mutants shown in Figure 3-17 (Section 3.4.3.2).

## 7.4 HRES *recq2* I-SceI assay PCRs

The results of the diagnostic PCRs to test for *PIF6* (as a positive control), *VSG221* and *ESAG1* in survivors from the first *RECQ2* HRES I-SceI assay (Section 4.5.1) are shown in Figure 7-2.

A summary of the analysis of survivors from the first *RECQ2* HRES I-SceI survival assay is shown in Figure 7-2. Clones determined to not be induced, due to puromycin resistance and presence of both *VSG221* and *ESAG1*. Several clones were possibly uninduced, due to puromycin resistance. Both clones classed as ‘uninduced’ and ‘possibly uninduced’ were excluded from calculation of the percentage of cells that retained *VSG221* and *ESAG1* (see Figure 4-12).





**Figure 7-2 Diagnostic PCRs from first *RECQ2* HRES I-SceI assay**  
 Genomic DNA was extracted from HRES wild type and HRES *recq2*<sup>-/-</sup> and HRES *recq2*<sup>-/-</sup> surviving clones from the first *RECQ2* HRES I-SceI assay (Section 4.5.1). The presence of genomic DNA was tested using PCR amplification of a region of *PIF6* using primers using primers #81 & #82 (245 bp product). Presence of *VSG221* was tested by PCR amplification using primers #90 and #91 (955 bp product) and presence of *ESAG1* was tested by PCR amplification using primers #92 and #93 (328 bp product). Distilled water (-) was used as a negative control. WT, wild type (positive control). -T, survivors from non-tetracycline (I-SceI non-induced) condition; +T, survivors from with tetracycline (I-SceI induced) condition. Numbers indicate clone numbers. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels/membranes or from disparate parts of the same gel/membrane. Sizes shown (ladder, bp).



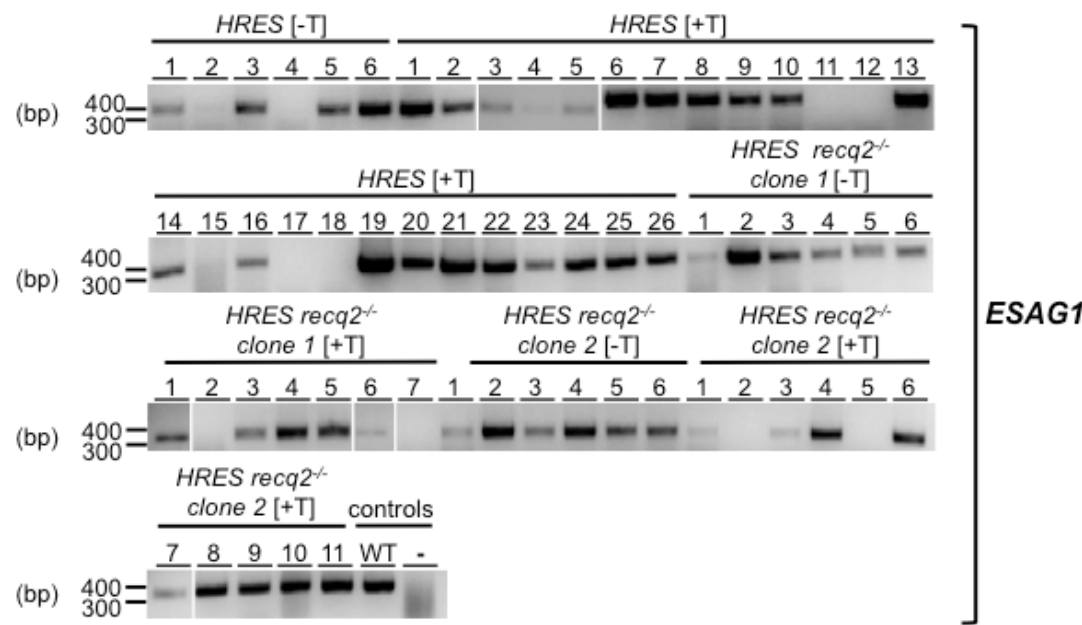


Figure 7-2 continued

| Cell line                               | Induction | Clone | Survival in PUR | PCR  |        |       |
|---|-----------|-------|-----------------|------|--------|-------|
|   |           |       |                 | PIF6 | VSG221 | ESAG1 |
| HRES                                    | -         | 1     |                 |      |        |       |
|   |           | 2     |                 |      |        |       |
|   |           | 3     |                 |      |        |       |
|   |           | 4     |                 |      |        |       |
|   |           | 5     |                 |      |        |       |
|   |           | 6     |                 |      |        |       |
|   | +         | 1     |                 |      |        |       |
|   |           | 2     |                 |      |        |       |
|   |           | 3     |                 |      |        |       |
|   |           | 4     |                 |      |        |       |
|   |           | 5     |                 |      |        |       |
|   |           | 6     |                 |      |        |       |
|   |           | 7     |                 |      |        |       |
|   |           | 8     |                 |      |        |       |
|   |           | 9     |                 |      |        |       |
|   |           | 10    |                 |      |        |       |
|   |           | 11    |                 |      |        |       |
|   |           | 12    |                 |      |        |       |
|   |           | 13    |                 |      |        |       |
|   |           | 14    |                 |      |        |       |
|   |           | 15    |                 |      |        |       |
|   |           | 16    |                 |      |        |       |
|   |           | 17    |                 |      |        |       |
|   |           | 18    |                 |      |        |       |
|   |           | 19    |                 |      |        |       |
|   |           | 20    |                 |      |        |       |
|   |           | 21    |                 |      |        |       |
|   |           | 22*   |                 |      |        |       |
|   |           | 23    |                 |      |        |       |
|   |           | 24    |                 |      |        |       |
|   |           | 25    |                 |      |        |       |
|   |           | 26    |                 |      |        |       |
| HRES <i>recq2</i> <sup>-/-</sup> Clone1 | -         | 1     |                 |      |        |       |
|   |           | 2     |                 |      |        |       |
|   |           | 3     |                 |      |        |       |
|   |           | 4     |                 |      |        |       |
|   |           | 5     |                 |      |        |       |
|   |           | 6     |                 |      |        |       |
|   |           | 1     |                 |      |        |       |
|   |           | 2     |                 |      |        |       |
|   |           | 3     |                 |      |        |       |
|   |           | 4*    |                 |      |        |       |
|   |           | 5     |                 |      |        |       |
|   |           | 6     |                 |      |        |       |
|   |           | 7     |                 |      |        |       |

Table 7-3      Summary of analysis of first *RECQ2* HRES I-SceI assay

Summary of survivors from first *RECQ2* HRES I-SceI assay (see Section 4.5 and Figure 7-2). Puromycin resistance was determined by testing growth in 1  $\mu\text{g.mL}^{-1}$  puromycin (PUR). Details of PCR analyses can be found in Figure 7-2 and Section 4.5.1. Grey shaded clones were not included in calculation of *VSG221* and *ESAG1* frequency (Fig. 4-12) due to: \*, clone not induced; §, clone likely not induced; #, no PCR data. Yellow, positive result; blue, negative result. Blue hatching, clones tested by PCR but no amplifiable genomic DNA (gDNA) recovered.

|   |   |     |  |  |  |  |
|---|---|-----|--|--|--|--|
| <i>HRES recq2</i> <sup>-/-</sup> Clone2 | - | 1   |  |  |  |  |
|   |   | 2   |  |  |  |  |
|   |   | 3   |  |  |  |  |
|   |   | 4   |  |  |  |  |
|   |   | 5   |  |  |  |  |
|   |   | 6   |  |  |  |  |
|   |   | 1   |  |  |  |  |
|   |   | 2*  |  |  |  |  |
|   |   | 3   |  |  |  |  |
|   |   | 4   |  |  |  |  |
|   |   | 5#  |  |  |  |  |
|   |   | 6*  |  |  |  |  |
|   |   | 7*  |  |  |  |  |
|   |   | 8*  |  |  |  |  |
|   |   | 9§  |  |  |  |  |
|   |   | 10§ |  |  |  |  |
|   |   | 11§ |  |  |  |  |
|   |   |     |  |  |  |  |
|   |   |     |  |  |  |  |
|   |   |     |  |  |  |  |

Negative

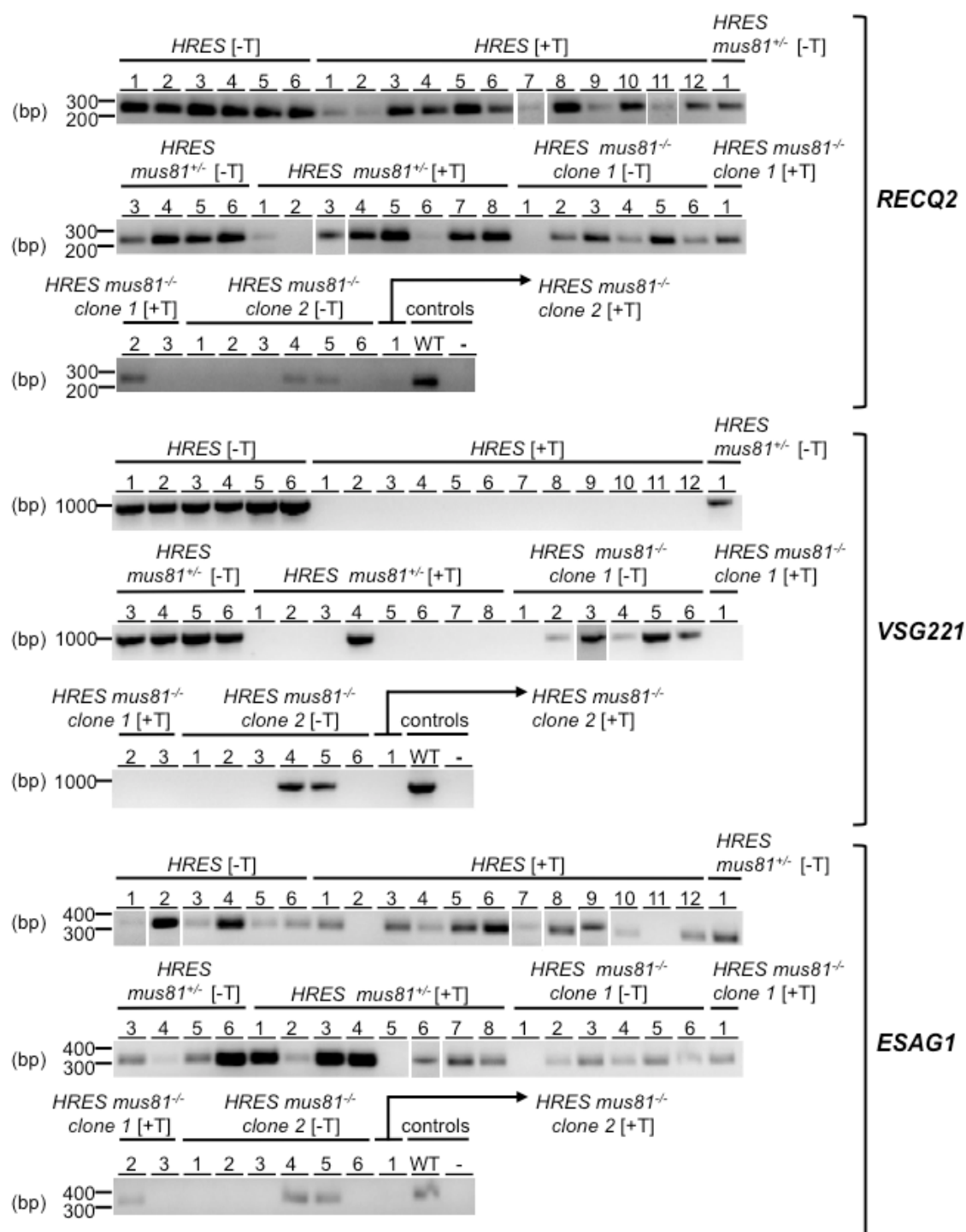
Positive

Clone tested, no amplifiable gDNA in sample

Table 7-3 continued

7.5 HRES *mus81* I-SceI assay PCRs

The results of the diagnostic PCRs to test for *RECQ2* (as a positive control), *VSG221* and *ESAG1* in survivors from the *MUS81* HRES I-SceI assay discussed in Section 4.5.2 are shown in Figure 7-3. Several clones determined to not be induced, due to puromycin resistance and presence of both *VSG221* and *ESAG1*, are marked in Figure 7-3. For several clones, genomic DNA was not recovered or cells not recovered for puromycin sensitivity - marked in Table 7-4. Clones classed as uninduced, or for which there was missing data were excluded from calculation of the percentage of cells that retained *VSG221* and *ESAG1* (see Figure 4-14).



**Figure 7-3** Diagnostic PCRs of survivors from *MUS81* HRES I-SceI assay. Genomic DNA was extracted from HRES wild type and HRES *mus81*<sup>-/-</sup> and HRES *mus81*<sup>-/-</sup> surviving clones from the *MUS81* HRES I-SceI assay detailed in Section 4.5.2. The presence of genomic DNA was tested using PCR amplification of a region of *RECQ2* using primers using primers #77 & #78 (232 bp product). Presence of *VSG221* was tested by PCR amplification using primers #90 and #91 (955 bp product) and presence of *ESAG1* was tested by PCR amplification using primers #92 and #93 (328 bp product). Distilled water (-) was used as a negative control. WT, wild type (positive control). -T, survivors from non-tetracycline (I-SceI non-induced) condition; +T, survivors from with tetracycline (I-SceI induced) condition. Numbers indicate clone numbers. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels/membranes or from disparate parts of the same gel/membrane. Sizes shown, (ladder, bp).

| Cell line                               | Induction | Clone          | Survival in PUR | PCR   |        |       |
|---|-----------|----------------|-----------------|-------|--------|-------|
|   |           |                |                 | RECQ2 | VSG221 | ESAG1 |
| HRES                                    | -         | 1              |                 |       |        |       |
|   |           | 2              |                 |       |        |       |
|   |           | 3              |                 |       |        |       |
|   |           | 4              |                 |       |        |       |
|   |           | 5              |                 |       |        |       |
|   |           | 6              |                 |       |        |       |
|   | +         | 1              |                 |       |        |       |
|   |           | 2 <sup>#</sup> |                 |       |        |       |
|   |           | 3              |                 |       |        |       |
|   |           | 4              |                 |       |        |       |
|   |           | 5              |                 |       |        |       |
|   |           | 6              |                 |       |        |       |
| HRES <i>mus81</i> <sup>+/-</sup>        | -         | 1              |                 |       |        |       |
|   |           | 2              |                 |       |        |       |
|   |           | 3              |                 |       |        |       |
|   |           | 4              |                 |       |        |       |
|   |           | 5              |                 |       |        |       |
|   |           | 6              |                 |       |        |       |
|   | +         | 1              |                 |       |        |       |
|   |           | 2              |                 |       |        |       |
| HRES <i>mus81</i> <sup>-/-</sup> Clone1 | -         | 1 <sup>#</sup> |                 |       |        |       |
|   |           | 2              |                 |       |        |       |
|   |           | 3              |                 |       |        |       |
|   |           | 4              |                 |       |        |       |
|   |           | 5              |                 |       |        |       |
|   |           | 6              |                 |       |        |       |
|   | +         | 1              |                 |       |        |       |
|   |           | 2              |                 |       |        |       |
|   |           | 3 <sup>#</sup> |                 |       |        |       |

**Table 7-4 Summary of analysis of first *MUS81* HRES I-SceI assay**  
 Summary of survivors from *MUS81* HRES I-SceI assay (see Section 4.5.2 and Figure 7-3). Puromycin resistance was determined by testing growth in 1  $\mu\text{g.mL}^{-1}$  puromycin (PUR). Details of PCR analyses can be found in Figure 7-3 and Section 4.5. Grey shaded clones were not included in calculation of VSG221 and ESAG1 frequency (Fig. 4-15c) due to: \*, clone not induced; #, no PCR data. Yellow, positive result; blue, negative result; blank box, cells not recovered for puromycin (PURO) testing; blue hatching, clones tested by PCR but not amplifiable genomic DNA (gDNA) recovered.

|   |   |                |  |  |  |  |
|---|---|----------------|--|--|--|--|
| HRES <i>mus81</i> <sup>-/-</sup><br>Clone 2 | - | 1              |  |  |  |  |
|   |   | 2 <sup>#</sup> |  |  |  |  |
|   |   | 3 <sup>#</sup> |  |  |  |  |
|   |   | 4 <sup>*</sup> |  |  |  |  |
|   |   | 5 <sup>*</sup> |  |  |  |  |
|   |   | 6              |  |  |  |  |
|   | + | 1 <sup>#</sup> |  |  |  |  |

Negative

Positive

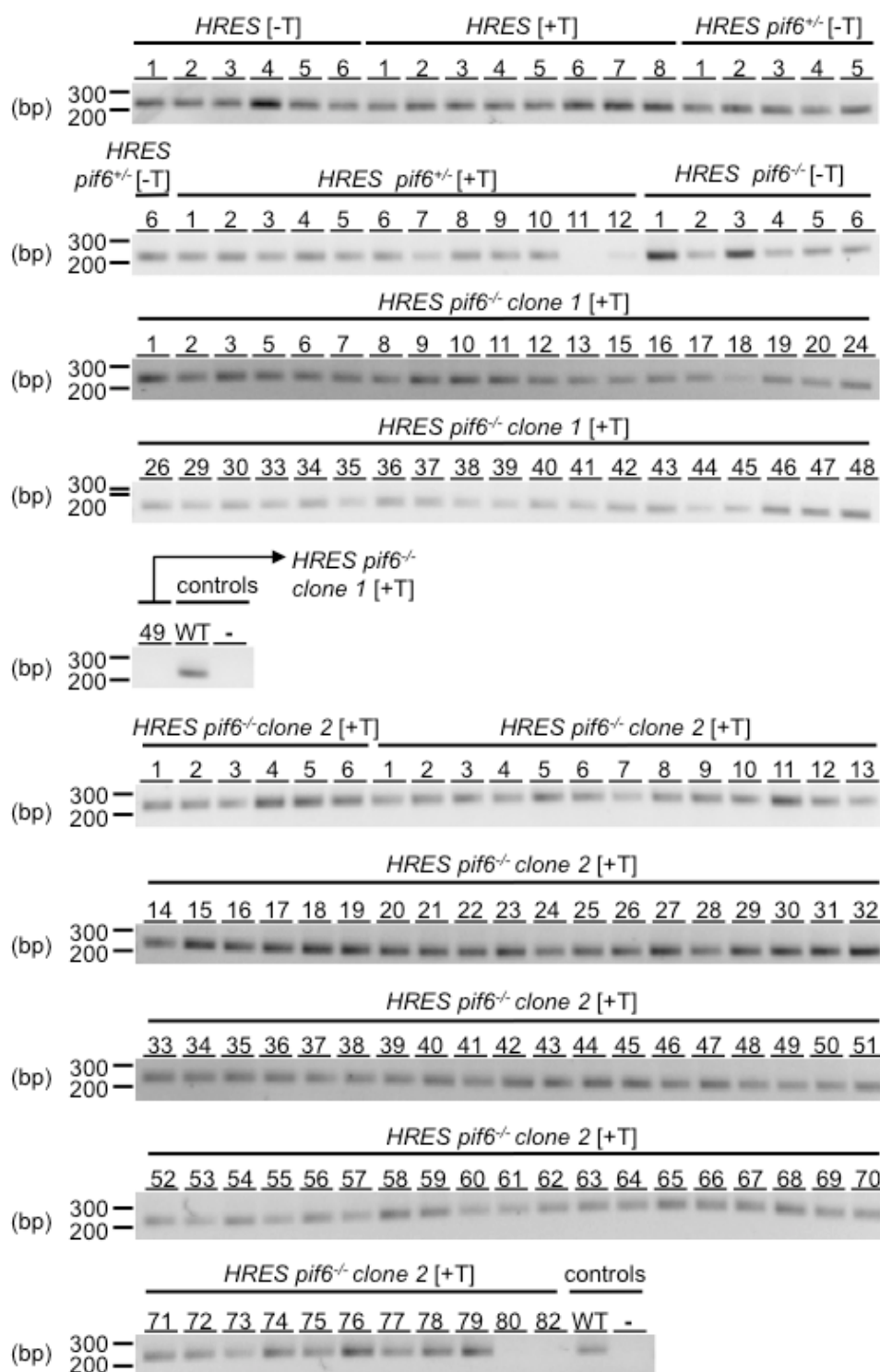
No growth with or without PURO

Clone tested, no amplifiable gDNA in sample

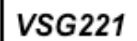
Table 7-4 continued

7.6 HRES *pif6* I-SceI assay PCR

The results of the diagnostic PCRs to test for *RECQ2* (as a positive control), *VSG221* and *ESAG1*, in survivors from the *MUS81* HRES I-SceI assay discussed in Section 4.5.3 are shown in Figure 7-4. Clones determined to not be induced, due to puromycin resistance and presence of both *VSG221* and *ESAG1*, are marked in Figure 7-4. For several clones, genomic DNA was not recovered or cells were not recovered for puromycin sensitivity testing - marked in Table 7-5. Clones classed as uninduced, or for which there was missing data (marked in Table 7-5) were excluded from calculation of the percentage of cells that retained *VSG221* and *ESAG1* (see Figure 4-15).



**Figure 7-4** Diagnostic PCRs of survivors from first *PIF6* HRES I-SceI assay  
Genomic DNA was extracted from HRES wild type and HRES *pfif6*<sup>+/-</sup> and HRES *pfif6*<sup>-/-</sup> surviving clones from the *PIF6* HRES I-SceI assay detailed in Section 4.5.3. The presence of genomic DNA was tested using PCR amplification of a region of *RECQ2* using primers using primers #77 & #78 (232 bp product). Presence of *VSG221* was tested by PCR amplification using primers #90 and #91 (955 bp product) and presence of *ESAG1* was tested by PCR amplification using primers #92 and #93 (328 bp product). Distilled water (-) was used as a negative control. WT, wild type (positive control). -T, survivors from non-tetracycline (I-SceI non-induced) condition; +T, survivors from with tetracycline (I-SceI induced) condition. Numbers indicate clone numbers. Gaps indicate that lanes have been aligned in this figure after excision from multiple gels/membranes or from disparate parts of the same gel/membrane. Sizes shown, (ladder, bp).



**Figure 7-4 continued**



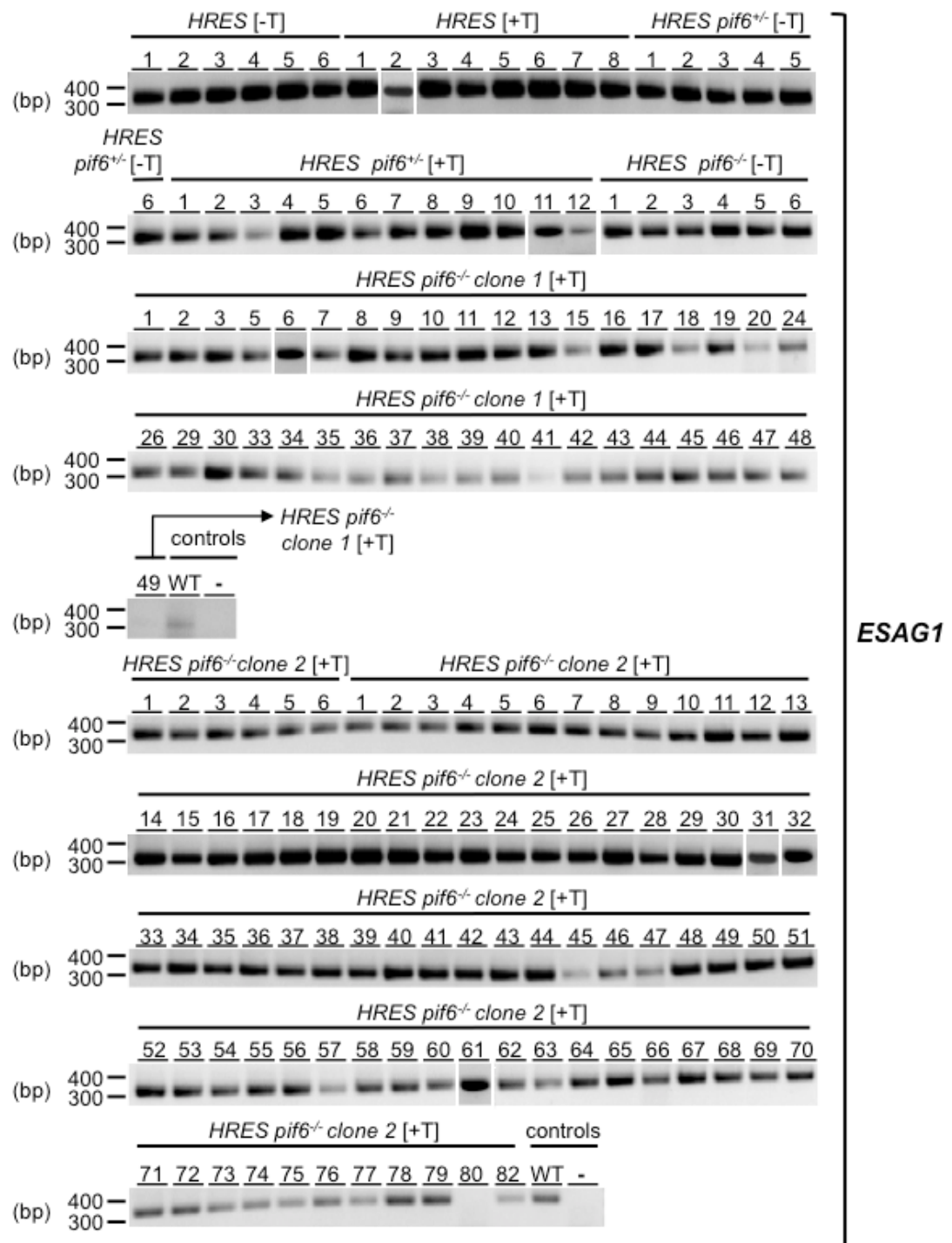


Figure 7-4 continued

| Cell line                     | Induction | Clone | Survival in PUR | PCR   |        |       |
|-------------------------------|-----------|-------|-----------------|-------|--------|-------|
|                               |           |       |                 | RECQ2 | VSG221 | ESAG1 |
| HRES                          | -         | 1     |                 |       |        |       |
|                               |           | 2     |                 |       |        |       |
|                               |           | 3     |                 |       |        |       |
|                               |           | 4     |                 |       |        |       |
|                               |           | 5     |                 |       |        |       |
|                               |           | 6     |                 |       |        |       |
|                               | +         | 1     |                 |       |        |       |
|                               |           | 2     |                 |       |        |       |
|                               |           | 3     |                 |       |        |       |
|                               |           | 4     |                 |       |        |       |
|                               |           | 5     |                 |       |        |       |
|                               |           | 6     |                 |       |        |       |
|                               |           | 7     |                 |       |        |       |
|                               |           | 8     |                 |       |        |       |
|                               |           | 9     |                 |       |        |       |
|                               |           | 10    |                 |       |        |       |
|                               |           | 11    |                 |       |        |       |
|                               |           | 12    |                 |       |        |       |
|                               |           | 13    |                 |       |        |       |
|                               |           | 14    |                 |       |        |       |
| <i>HRES</i><br><i>pif6+/-</i> | -         | 1     |                 |       |        |       |
|                               |           | 2     |                 |       |        |       |
|                               |           | 3     |                 |       |        |       |
|                               |           | 4     |                 |       |        |       |
|                               |           | 5     |                 |       |        |       |
|                               |           | 6     |                 |       |        |       |
|                               | +         | 1     |                 |       |        |       |
|                               |           | 2     |                 |       |        |       |
|                               |           | 3     |                 |       |        |       |
|                               |           | 4     |                 |       |        |       |
|                               |           | 5     |                 |       |        |       |
|                               |           | 6     |                 |       |        |       |
|                               |           | 7     |                 |       |        |       |
|                               |           | 8     |                 |       |        |       |
|                               |           | 9     |                 |       |        |       |
|                               |           | 10    |                 |       |        |       |
|                               |           | 11    |                 |       |        |       |
|                               |           | 12    |                 |       |        |       |

**Table 7-5** Summary of analysis of survivors from first *PIF6* HRES I-SceI assay  
Summary of survivors from *MUS81* HRES I-SceI assay (see Section 4.5.3 and Figure 7-4).  
Puromycin resistance was determined by testing growth in 1  $\mu\text{g.mL}^{-1}$  puromycin (PUR).  
Details of PCR analyses can be found in Figure 7-3 and Section 4.5. Grey shaded clones  
were not included in calculation of *VSG221* and *ESAG1* frequency (Fig. 4-16C) due to \*,  
clone not induced; #, no PCR data. Yellow, positive result; blue, negative result; blank box,  
cells not recovered for puromycin (PURO) testing; blue hatching, clones tested by PCR but  
not amplifiable genomic DNA (gDNA) recovered; pink hatching, genomic DNA not obtained.

|   |   |                 |  |  |  |  |
|---|---|-----------------|--|--|--|--|
| <i>HRES</i><br><i>pif6</i> <sup>-/-</sup><br>Clone1 | - | 1               |  |  |  |  |
|   |   | 2               |  |  |  |  |
|   |   | 3               |  |  |  |  |
|   |   | 4               |  |  |  |  |
|   |   | 5               |  |  |  |  |
|   |   | 6               |  |  |  |  |
|   | + | 1               |  |  |  |  |
|   |   | 2               |  |  |  |  |
|   |   | 3               |  |  |  |  |
|   |   | 4               |  |  |  |  |
|   |   | 5               |  |  |  |  |
|   |   | 6               |  |  |  |  |
|   |   | 7               |  |  |  |  |
|   |   | 8               |  |  |  |  |
|   |   | 9               |  |  |  |  |
|   |   | 10              |  |  |  |  |
|   |   | 11              |  |  |  |  |
|   |   | 12              |  |  |  |  |
|   |   | 13              |  |  |  |  |
|   |   | 14 <sup>#</sup> |  |  |  |  |
|   |   | 15              |  |  |  |  |
|   |   | 16              |  |  |  |  |
|   |   | 17              |  |  |  |  |
|   |   | 18              |  |  |  |  |
|   |   | 19              |  |  |  |  |
|   |   | 20              |  |  |  |  |
|   |   | 21 <sup>#</sup> |  |  |  |  |
|   |   | 22 <sup>#</sup> |  |  |  |  |
|   |   | 23 <sup>#</sup> |  |  |  |  |
|   |   | 24              |  |  |  |  |
|   |   | 25 <sup>#</sup> |  |  |  |  |
|   |   | 26              |  |  |  |  |
|   |   | 27 <sup>#</sup> |  |  |  |  |
|   |   | 28 <sup>#</sup> |  |  |  |  |
|   |   | 29              |  |  |  |  |
|   |   | 30              |  |  |  |  |
|   |   | 31 <sup>#</sup> |  |  |  |  |
|   |   | 32 <sup>#</sup> |  |  |  |  |
|   |   | 33              |  |  |  |  |
|   |   | 34              |  |  |  |  |
|   |   | 35              |  |  |  |  |
|   |   | 36              |  |  |  |  |
|   |   | 37              |  |  |  |  |
|   |   | 38              |  |  |  |  |
|   |   | 39              |  |  |  |  |
|   |   | 40*             |  |  |  |  |

Table 7-5 continued

|   |   |     |  |  |  |  |
|---|---|-----|--|--|--|--|
|   |   | 41  |  |  |  |  |
|   |   | 42  |  |  |  |  |
|   |   | 43  |  |  |  |  |
|   |   | 44  |  |  |  |  |
|   |   | 45  |  |  |  |  |
|   |   | 46  |  |  |  |  |
|   |   | 47  |  |  |  |  |
|   |   | 48* |  |  |  |  |
|   |   | 49# |  |  |  |  |
| <i>HRES</i><br><i>pif6</i> <sup>-/-</sup><br>Clone2 | - | 1   |  |  |  |  |
|   |   | 2   |  |  |  |  |
|   |   | 3   |  |  |  |  |
|   |   | 4   |  |  |  |  |
|   |   | 5   |  |  |  |  |
|   |   | 6   |  |  |  |  |
|   | + | 1   |  |  |  |  |
|   |   | 2   |  |  |  |  |
|   |   | 3   |  |  |  |  |
|   |   | 4   |  |  |  |  |
|   |   | 5   |  |  |  |  |
|   |   | 6   |  |  |  |  |
|   |   | 7   |  |  |  |  |
|   |   | 8   |  |  |  |  |
|   |   | 9   |  |  |  |  |
|   |   | 10  |  |  |  |  |
|   |   | 11  |  |  |  |  |
|   |   | 12  |  |  |  |  |
|   |   | 13  |  |  |  |  |
|   |   | 14  |  |  |  |  |
|   |   | 15  |  |  |  |  |
|   |   | 16  |  |  |  |  |
|   |   | 17  |  |  |  |  |
|   |   | 18  |  |  |  |  |
|   |   | 19  |  |  |  |  |
|   |   | 20  |  |  |  |  |
|   |   | 21  |  |  |  |  |
|   |   | 22  |  |  |  |  |
|   |   | 23  |  |  |  |  |
|   |   | 24  |  |  |  |  |
|   |   | 25  |  |  |  |  |
|   |   | 26  |  |  |  |  |
|   |   | 27  |  |  |  |  |
|   |   | 28  |  |  |  |  |
|   |   | 29  |  |  |  |  |
|   |   | 30  |  |  |  |  |
|   |   | 31  |  |  |  |  |
|   |   | 32  |  |  |  |  |
|   |   | 33  |  |  |  |  |

Table 7-5 continued

|  |  |     |    |  |  |  |
|--|--|-----|----|--|--|--|
|  |  |     | 34 |  |  |  |
|  |  |     | 35 |  |  |  |
|  |  |     | 36 |  |  |  |
|  |  |     | 37 |  |  |  |
|  |  |     | 38 |  |  |  |
|  |  |     | 39 |  |  |  |
|  |  |     | 40 |  |  |  |
|  |  |     | 41 |  |  |  |
|  |  |     | 42 |  |  |  |
|  |  |     | 43 |  |  |  |
|  |  |     | 44 |  |  |  |
|  |  |     | 45 |  |  |  |
|  |  |     | 46 |  |  |  |
|  |  |     | 47 |  |  |  |
|  |  |     | 48 |  |  |  |
|  |  |     | 49 |  |  |  |
|  |  |     | 50 |  |  |  |
|  |  | 51* |    |  |  |  |
|  |  |     | 52 |  |  |  |
|  |  |     | 53 |  |  |  |
|  |  |     | 54 |  |  |  |
|  |  |     | 55 |  |  |  |
|  |  |     | 56 |  |  |  |
|  |  |     | 57 |  |  |  |
|  |  |     | 58 |  |  |  |
|  |  |     | 59 |  |  |  |
|  |  |     | 60 |  |  |  |
|  |  |     | 61 |  |  |  |
|  |  |     | 62 |  |  |  |
|  |  |     | 63 |  |  |  |
|  |  |     | 64 |  |  |  |
|  |  |     | 65 |  |  |  |
|  |  |     | 66 |  |  |  |
|  |  |     | 67 |  |  |  |
|  |  |     | 68 |  |  |  |
|  |  |     | 69 |  |  |  |
|  |  |     | 70 |  |  |  |
|  |  |     | 71 |  |  |  |
|  |  |     | 72 |  |  |  |
|  |  |     | 73 |  |  |  |
|  |  |     | 74 |  |  |  |
|  |  |     | 75 |  |  |  |
|  |  |     | 76 |  |  |  |
|  |  |     | 77 |  |  |  |
|  |  | 78* |    |  |  |  |
|  |  |     | 79 |  |  |  |
|  |  | 80# |    |  |  |  |

Table 7-5 continued

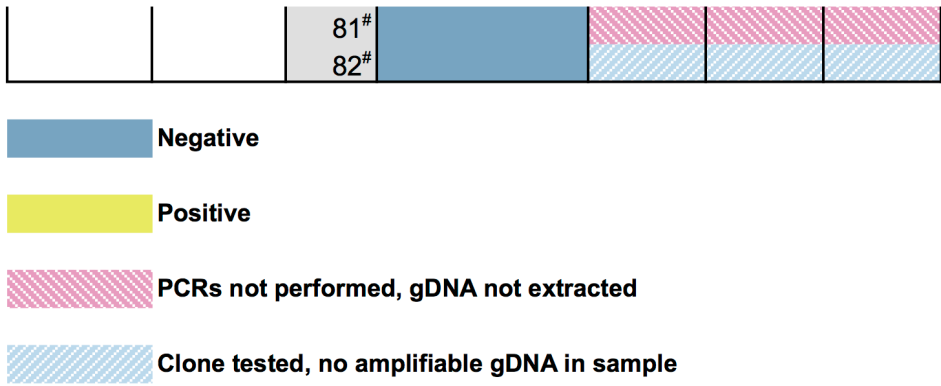


Table 7-5 continued

## List of References

- Abe T., Yoshimura A., Hosono Y., Tada S., Seki M., Enomoto T. (2011) The N-terminal region of RECQL4 lacking the helicase domain is both essential and sufficient for the viability of vertebrate cells: Role of the N-terminal region of RECQL4 in cells. *Biochimica et Biophysica Acta* **1813**, 473-479
- Adams M. D., McVey M., Sekelsky J. J. (2003) *Drosophila* BLM in double-strand break repair by synthesis-dependent strand annealing. *Science* **299**, 265-267
- Aitcheson N., Talbot S., Shapiro J., Hughes K., Adkin C., Butt T., Sheader K., Rudenko G. (2005) VSG switching in *Trypanosoma brucei*: antigenic variation analysed using RNAi in the absence of immune selection. *Molecular microbiology* **57**, 1608-1622
- Albino A. P., Huang X., Jorgensen E. D., Gietl D., Traganos F., Darzynkiewicz Z. (2006) Induction of DNA double-strand breaks in A549 and normal human pulmonary epithelial cells by cigarette smoke is mediated by free radicals. *International journal of oncology* **28**, 1491-1505
- Alibu V. P., Storm L., Haile S., Clayton C., Horn D. (2005) A doubly inducible system for RNA interference and rapid RNAi plasmid construction in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **139**, 75-82
- Aline R., Macdonald G., Brown E., Allison J., Myler P., Rothwell V., Stuart K. (1985) (TAA)<sub>n</sub> within sequences flanking several intrachromosomal variant surface glycoproteins in *Trypanosoma brucei*. *Nucleic Acids Research* **13**, 3161-3177
- Alsford S., Glover L., Horn D. (2005a) Multiplex analysis of RNA interference defects in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **139**, 129-132
- Alsford S., Horn D. (2008) Single-locus targeting constructs for reliable regulated RNAi and transgene expression in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **161**, 76-79
- Alsford S., Horn D., Glover L. (2009) DNA breaks as triggers for antigenic variation in African trypanosomes. *Genome Biology* **10**
- Alsford S., Kawahara T., Glover L., Horn D. (2005b) Tagging a *T. brucei* RRNA locus improves stable transfection efficiency and circumvents inducible expression position effects. *Molecular and Biochemical Parasitology* **144**, 142-148
- Anand R. P., Shah K. A., Niu H. Y., Sung P., Mirkin S. M., Freudenreich C. H. (2012) Overcoming natural replication barriers: differential helicase requirements. *Nucleic Acids Research* **40**, 1091-1105
- Aravind L., Koonin E. V. (2000) SAP - a putative DNA-binding motif involved in chromosomal organization. *Trends in Biochemical Sciences* **25**, 112-114

- Arazoe T., Younomaru T., Ohsato S., Kimura M., Arie T., Kuwata S. (2014) Site-specific DNA double-strand break generated by I-SceI endonuclease enhances ectopic homologous recombination in *Pyricularia oryzae*. *FEMS Microbiology Letters* **352**, 221-229
- Argueso J. L., Kijas A. W., Sarin S., Heck J., Waase M., Alani E. (2003) Systematic mutagenesis of the *Saccharomyces cerevisiae* MLH1 gene reveals distinct roles for Mlh1p in meiotic crossing over and in vegetative and meiotic mismatch repair. *Molecular and Cellular Biology* **23**, 873-886
- Azvolinsky A., Dunaway S., Torres J. Z., Bessler J. B., Zakian V. A. (2006) The *S. cerevisiae* Rrm3p DNA helicase moves with the replication fork and affects replication of all yeast chromosomes. *Genes & Development* **20**, 3104-3116
- Bachrati C. Z., Hickson I. D. (2003) RecQ helicases: suppressors of tumorigenesis and premature aging. *Biochemical Journal* **374**, 577-606
- Bachrati C. Z., Hickson I. D. (2008) RecQ helicases: guardian angels of the DNA replication fork. *Chromosoma* **117**, 219-233
- Barbet A. F., Davis W. C., McGuire T. C. (1982) Cross-neutralization of two different trypanosome populations derived from a single organism. *Nature* **300**, 453-456
- Barbet A. F., Myler P. J., Williams R. O., McGuire T. C. (1989) Shared surface epitopes among trypanosomes of the same serodeme expressing different variable surface glycoprotein genes. *Molecular and Biochemical Parasitology* **32**, 191-199
- Barnes R. L., McCulloch R. (2007) Trypanosoma brucei homologous recombination is dependent on substrate length and homology, though displays a differential dependence on mismatch repair as substrate length decreases. *Nucleic Acids Research* **35**, 3478-3493
- Barnes R. L., Shi H., Kolev N. G., Tschudi C., Ullu E. (2012) Comparative genomics reveals two novel rai factors in Trypanosoma brucei and provides insight into the core machinery. *PLoS Pathogens* **8**
- Barrett M. P., Burchmore R. J. S., Stich A., Lazzari J. O., Frasch A. C., Cazzulo J. J., Krishna S. (2003) The Trypanosomiasis. *The Lancet* **362**, 1469-1480
- Barry D., McCulloch R. (2009) Molecular microbiology: A key event in survival. *Nature* **459**, 172-173
- Barry J. D. (1997) The relative significance of mechanisms of antigenic variation in African trypanosomes. *Parasitology Today* **13**, 212-218
- Barry J. D., Marcello L., Morrison L. J., Read A. F., Lythgoe K., Jones N., Carrington M., Blandin G., Bohme U., Caler E., Hertz-Fowler C., Renaud H., El-Sayed N., Berriman M. (2005) What the genome sequence



is revealing about trypanosome antigenic variation. *Biochemical Society Transactions* **33**, 986-989

Barry J. D., McCulloch R. (2001) Antigenic variation in trypanosomes: Enhanced phenotypic variation in a eukaryotic parasite. *Advances in Parasitology* **49**, 1-70

Bastin-Shanower S. A., Fricke W. M., Mullen J. R., Brill S. J. (2003) The mechanism of Mus81-Mms4 cleavage site selection distinguishes it from the homologous endonuclease Rad1-Rad10. *Molecular and Cellular Biology* **23**, 3487-3496

Bazzi M., Mantiero D., Trovesi C., Lucchini G., Longhese M. P. (2010) Dephosphorylation of gamma H2A by Glc7/Protein Phosphatase 1 promotes recovery from inhibition of DNA replication. *Molecular and Cellular Biology* **30**, 131-145

Becker M., Aitcheson N., Byles E., Wickstead B., Louis E., Rudenko G. (2004) Isolation of the repertoire of VSG expression site containing telomeres of *Trypanosoma brucei* 427 using transformation-associated recombination in yeast. *Genome Research* **14**, 2319-2329

Bell J. S., McCulloch R. (2003) Mismatch repair regulates homologous recombination, but has little influence on antigenic variation, in *Trypanosoma brucei*. *Journal of Biological Chemistry* **278**, 45182-45188

Benmerzouga I., Concepción-Acevedo J., Kim H.-S., Vandomos A. V., Cross G. A. M., Klingbeil M. M., Li B. (2013) *Trypanosoma brucei* Orc1 is essential for nuclear DNA replication and affects both VSG silencing and VSG switching. *Molecular Microbiology* **87**, 196-210

Bergink S., Ammon T., Kern M., Schermelleh L., Leonhardt H., Jentsch S. (2013) Role of Cdc48/p97 as a SUMO-targeted segregase curbing Rad51-Rad52 interaction. *Nature Cell Biology* **15**, 526-532

Bernards A., Vanderploeg L. H. T., Frasch A. C. C., Borst P., Boothroyd J. C., Coleman S., Cross G. A. M. (1981) Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3' end. *Cell* **27**, 497-505

Bernstein K. A., Gangloff S., Rothstein R. (2010) The RecQ DNA helicases in DNA repair. *Annual Review of Genetics* **44**, 393-417.

Bernstein K. A., Shor E., Sunjevaric I., Fumasoni M., Burgess R. C., Foiani M., Branzei D., Rothstein R. (2009) Sgs1 function in the repair of DNA replication intermediates is separable from its role in homologous recombinational repair. *EMBO Journal* **28**, 915-925

Berriman M., Ghedin E., Hertz-Fowler C., Blandin G., Renauld H., Bartholomeu D. C., Lennard N. J., Caler E., Hamlin N. E., Haas B., Bohme W., Hannick L., Aslett M. A., Shallom J., Marcello L., Hou L. H., Wickstead B., Alsmark U. C. M., Arrowsmith C., Atkin R. J., Barron A. J., Bringaud F., Brooks K., Carrington M., Cherevach I., Chillingworth

- T. J., Churcher C., Clark L. N., Corton C. H., Cronin A., Davies R. M., Doggett J., Djikeng A., Feldblyum T., Field M. C., Fraser A., Goodhead I., Hance Z., Harper D., Harris B. R., Hauser H., Hostetter J., Ivens A., Jagels K., Johnson D., Johnson J., Jones K., Kerhornou A. X., Koo H., Larke N., Landfear S., Larkin C., Leech V., Line A., Lord A., MacLeod A., Mooney P. J., Moule S., Martin D. M. A., Morgan G. W., Mungall K., Norbertczak H., Ormond D., Pai G., Peacock C. S., Peterson J., Quail M. A., Rabinowitsch E., Rajandream M. A., Reitter C., Salzberg S. L., Sanders M., Schobel S., Sharp S., Simmonds M., Simpson A. J., Talton L., Turner C. M. R., Tait A., Tivey A. R., Van Aken S., Walker D., Wanless D., Wang S. L., White B., White O., Whitehead S., Woodward J., Wortman J., Adams M. D., Embley T. M., Gull K., Ullu E., Barry J. D., Fairlamb A. H., Opperdoes F., Barret B. G., Donelson J. E., Hall N., Fraser C. M., Melville S. E., El-Sayed N. M. (2005) The genome of the African trypanosome *Trypanosoma brucei*. *Science* **309**, 416-422
- Berti M., Chaudhuri A. R., Thangavel S., Gomathinayagam S., Kenig S., Vujanovic M., Odreman F., Glatter T., Graziano S., Mendoza-Maldonado R., Marino F., Lucic B., Biasin V., Gstaiger M., Aebersold R., Sidorova J. M., Monnat R. J., Lopes M., Vindigni A. (2013) Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition. *Nature Structural & Molecular Biology* **20**, 347-354
- Bianchi V., Pontis E., Reichard P. (1986) Changes of deoxyribonucleoside triphosphate pools induced by hydroxyurea and their relation to DNA-synthesis. *Journal of Biological Chemistry* **261**, 6037-6042
- Bitter W., Gerrits H., Kieft R., Borst P. (1998) The role of transferrin-receptor variation in the host range of *Trypanosoma brucei*. *Nature* **391**, 499-502
- Blanco M. G., Matos J., Rass U., Ip S. C. Y., West S. C. (2010) Functional overlap between the structure-specific nucleases Yen1 and Mus81-Mms4 for DNA-damage repair in *S. cerevisiae*. *DNA Repair* **9**, 394-402
- Bochman M. L., Paeschke K., Zakian V. A. (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nature Reviews Genetics* **13**, 770-780
- Bochman M. L., Sabouri N., Zakian V. A. (2010) Unwinding the functions of the Pif1 family helicases. *DNA Repair* **9**, 237-249
- Boddy M. N., Gaillard P. H. L., McDonald W. H., Shanahan P., Yates J. R., Russell P. (2001) Mus81-Eme1 are essential components of a Holliday junction resolvase. *Cell* **107**, 537-548
- Boddy M. N., Lopez-Girona A., Shanahan P., Interthal H., Heyer W. D., Russell P. (2000) Damage tolerance protein Mus81 associates with the FHA1 domain of checkpoint kinase Cds1. *Molecular and Cellular Biology* **20**, 8758-8766
- Boothroyd C. E., Dreesen O., Leonova T., Ly K. I., Figueiredo L. M., Cross G. A. M., Papavasiliou F. N. (2009) A yeast-endonuclease-generated DNA

- break induces antigenic switching in *Trypanosoma brucei*. *Nature* **459**, 278-281
- Borst P., Rudenko G., Taylor M. C., Blundell P. A., van Leeuwen F., Bitter W., Cross M., McCulloch R. (1996) Antigenic variation in trypanosomes. *Archives of Medical Research* **27**, 379-388
- Branzei D., Foiani M. (2008) Regulation of DNA repair throughout the cell cycle. *Nature Reviews Molecular Cell Biology* **9**, 297-308
- Bressan D. A., Baxter B. K., Petrini J. H. J. (1999) The Mre11-Rad50-Xrs2 protein complex facilitates homologous recombination-based double-strand break repair in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **19**, 7681-7687
- Brewer B. J. (1988) When polymerases collide - replication and the transcriptional organization of the *Escherichia coli* chromosome. *Cell* **53**, 679-686
- Brookes P., Lawley P. D. (1961) Reaction of mono- and di-functional alkylating agents with nucleic acids. *Biochemical Journal* **80**, 496-503
- Bruhn D. F., Mozeleski B., Falkin L., Klingbeil M. M. (2010) Mitochondrial DNA polymerase POLIB is essential for minicircle DNA replication in African trypanosomes. *Molecular Microbiology* **75**, 1414-1425
- Budd M. E., Reis C. C., Smith S., Myung K., Campbell J. L. (2006) Evidence suggesting that Pif1 helicase functions in DNA replication with the DNA2 helicase/nuclease and DNA polymerase delta. *Molecular and Cellular Biology* **26**, 2490-2500
- Burrow A. A., Marullo A., Holder L. R., Wang Y.-H. (2010) Secondary structure formation and DNA instability at fragile site FRA16B. *Nucleic Acids Research* **38**, 2865-2877
- Burton P., McBride D. J., Wilkes J. M., Barry J. D., McCulloch R. (2007) Ku heterodimer-independent end joining in *Trypanosoma brucei* cell extracts relies upon sequence microhomology. *Eukaryotic Cell* **6**, 1773-1781
- Bussen W., Raynard S., Busygina V., Singh A. K., Sung P. (2007) Holliday junction processing activity of the BLM-topo III alpha-BLAP75 complex. *Journal of Biological Chemistry* **282**, 31484-31492
- Cahoon L. A., Seifert H. S. (2009) An alternative DNA structure is necessary for pilin antigenic variation in *Neisseria gonorrhoeae*. *Science* **325**, 764-767
- Cahoon L. A., Seifert H. S. (2011) Focusing homologous recombination: pilin antigenic variation in the pathogenic *Neisseria*. *Molecular Microbiology* **81**, 1136-1143
- Cahoon L. A., Seifert H. S. (2013) Transcription of a cis-acting, noncoding, small RNA is required for pilin antigenic variation in *Neisseria gonorrhoeae*. *PLoS Pathogens* **9**

- Calderano S. G., Drosopoulos W. C., Quaresma M. M., Marques C. A., Kosiyatrakul S., McCulloch R., Schildkraut C. L., Elias M. C. (2015) Single molecule analysis of *Trypanosoma brucei* DNA replication dynamics. *Nucleic Acids Research* **43**, 2655-2665
- Calzada A., Hodgson B., Kanemaki M., Bueno A., Labib K. (2005) Molecular anatomy and regulation of a stable replisome eukaryotic DNA at a paused replication fork. *Genes & Development* **19**, 1905-1919
- Campbell D. A., Vanbree M. P., Boothroyd J. C. (1984) The 5' limit of transposition and upstream barren region of a trypanosome VSG gene tandem 76 base pair repeats flanking (TAA)<sup>90</sup>. *Nucleic Acids Research* **12**, 2759-2774
- Canman C. E. (2003) Checkpoint mediators: Relaying signals from DNA strand breaks. *Current Biology* **13**, R488-R490
- Cannavo E., Cejka P., Kowalczykowski S. C. (2013) Relationship of DNA degradation by *Saccharomyces cerevisiae* exonuclease 1 and its stimulation by RPA and Mre11-Rad50-Xrs2 to DNA end resection. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E1661-E1668
- Capp C., Wu J., Hsieh T.-s. (2009) *Drosophila* RecQ4 Has a 3'-5' DNA helicase activity that is essential for viability. *Journal of Biological Chemistry* **284**, 30845-30852
- Cejka P., Cannavo E., Polaczek P., Masuda-Sasa T., Pokharel S., Campbell J. L., Kowalczykowski S. C. (2010) DNA end resection by Dna2-Sgs1-RPA and its stimulation by Top3-Rmi1 and Mre11-Rad50-Xrs2. *Nature* **467**, 112-U149
- Chakraverty R. K., Hickson I. D. (1999) Defending genome integrity during DNA replication: a proposed role for RecQ family helicases. *Bioessays* **21**, 286-294
- Chandler J., Vondoros A. V., Mozeleski B., Klingbeil M. M. (2008) Stem-loop silencing reveals that a third mitochondrial DNA polymerase, POLID, is required for kinetoplast DNA replication in trypanosomes. *Eukaryotic Cell* **7**, 2141-2146
- Chang J. H., Kim J. J., Choi J. M., Lee J. H., Cho Y. (2008) Crystal structure of the Mus81-Eme1 complex. *Genes & Development* **22**, 1093-1106
- Chang M., Bellaoui M., Boone C., Brown G. W. (2002) A genome-wide screen for methyl methanesulfonate-sensitive mutants reveals genes required for S phase progression in the presence of DNA damage. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 16934-16939
- Chang M., Bellaoui M., Zhang C., Desai R., Morozov P., Delgado-Cruzata L., Rothstein R., Freyer G. A., Boone C., Brown G. W. (2005) RMI1/NCE4, a

- suppressor of genome instability, encodes a member of the RecQ helicase/Topo III complex. *EMBO Journal* **24**, 2024-2033
- Chapman J. R., Taylor Martin R. G., Boulton Simon J. (2012) Playing the end game: DNA double-strand break repair pathway choice. *Molecular Cell* **47**, 497-510
- Chaves I., Rudenko G., Dirks-Mulder A., Cross M., Borst P. (1999) Control of variant surface glycoprotein gene-expression sites in *Trypanosoma brucei*. *EMBO Journal* **18**, 4846-4855
- Chelysheva L., Vezon D., Belcram K., Gendrot G., Grelon M. (2008) The *Arabidopsis* BLAP75/Rmi1 homologue plays crucial roles in meiotic double-strand break repair. *PLoS Genetics* **4**
- Chen J. M., Cooper D. N., Chuzhanova N., Ferec C., Patrinos G. P. (2007) Gene conversion: mechanisms, evolution and human disease. *Nature Reviews Genetics* **8**, 762-775
- Chen L., Trujillo K., Ramos W., Sung P., Tomkinson A. E. (2001a) Promotion of Dnl4-catalyzed DNA end-joining by the Rad50/Mre11/Xrs2 and Hdf1/Hdf2 complexes. *Molecular Cell* **8**, 1105-1115
- Chen L., Trujillo K., Sung P., Tomkinson A. E. (2000) Interactions of the DNA ligase IV-XRCC4 complex with DNA ends and the DNA-dependent protein kinase. *Journal of Biological Chemistry* **275**, 26196-26205
- Chen X. B., Melchionna R., Denis C. M., Gaillard P. H. L., Blasina A., Van de Weyer I., Boddy M. N., Russell P., Vialard J., McGowan C. H. (2001b) Human Mus81-associated endonuclease cleaves Holliday junctions *in vitro*. *Molecular Cell* **8**, 1117-1127
- Chen Y. X., Dui W., Yu Z. S., Li C. Q., Ma J., Jiao R. J. (2010) *Drosophila* RecQ5 is required for efficient SSA repair and suppression of LOH *in vivo*. *Protein & Cell* **1**, 478-490
- Cheng X., Dunaway S., Ivessa A. S. (2007) The role of Pif1p, a DNA helicase in *Saccharomyces cerevisiae*, in maintaining mitochondrial DNA. *Mitochondrion* **7**, 211-222
- Cheng X., Qin Y., Ivessa A. S. (2009) Loss of mitochondrial DNA under genotoxic stress conditions in the absence of the yeast DNA helicase Pif1p occurs independently of the DNA helicase Rrm3p. *Molecular Genetics and Genomics* **281**, 635-645
- Chiolo I., Carotenuto W., Maffioletti G., Petrini J. H. J., Foiani M., Liberi G. (2005) Srs2 and Sgs1 DNA helicases associate with mre11 in different subcomplexes following checkpoint activation and CDK1-mediated Srs2 phosphorylation. *Molecular and Cellular Biology* **25**, 5738-5751
- Chung W. H., Zhu Z., Papusha A., Malkova A., Ira G. (2010) Defective resection at DNA double-strand breaks leads to *de novo* telomere formation and enhances gene targeting. *PLoS Genetics* **6**

- Churikov D., Price C. M. (2001) Telomeric and subtelomeric repeat sequences. In *eLS*. John Wiley & Sons, Ltd
- Ciccia A., Constantinou A., West S. C. (2003) Identification and characterization of the human Mus81-Eme1 endonuclease. *Journal of Biological Chemistry* **278**, 25172-25178
- Ciccia A., McDonald N., West S. C. (2008) Structural and functional relationships of the XPF/MUS81 family of proteins. *Annual Review of Biochemistry* **77**, 259-287
- Claessens A., Hamilton W. L., Kekre M., Otto T. D., Faizullahoy A., Rayner J. C., Kwiatkowski D. (2014) Generation of antigenic diversity in *Plasmodium falciparum* by structured rearrangement of *var* genes during mitosis. *PLoS Genetics* **10**
- Cobb J. A., Bjergbaek L. (2006) RecQ helicases: lessons from model organisms. *Nucleic Acids Research* **34**, 4106-4114
- Cobb J. A., Bjergbaek L., Shimada K., Frei C., Gasser S. M. (2003) DNA polymerase stabilization at stalled replication forks requires Mec1 and the RecQ helicase Sgs1. *EMBO Journal* **22**, 4325-4336
- Coin F., Auriol J., Tapias A., Clivio P., Vermeulen W., Egly J.-M. (2004) Phosphorylation of XPB helicase regulates TFIIH nucleotide excision repair activity. *EMBO Journal* **23**, 4835-4846
- Colleaux L., Dauriol L., Galibert F., Dujon B. (1988) Recognition and cleavage site of the intron-encoded omega transposase. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 6022-6026
- Concepcion-Acevedo J., Luo J., Klingbeil M. M. (2012) Dynamic localization of *Trypanosoma brucei* mitochondrial DNA polymerase ID. *Eukaryotic Cell* **11**, 844-855
- Conway C., McCulloch R., Ginger M. L., Robinson N. P., Browitt A., Barry J. D. (2002a) Ku is important for telomere maintenance, but not for differential expression of telomeric VSG genes, in African trypanosomes. *Journal of Biological Chemistry* **277**, 21269-21277
- Conway C., Proudfoot C., Burton P., Barry J. D., McCulloch R. (2002b) Two pathways of homologous recombination in *Trypanosoma brucei*. *Molecular Microbiology* **45**, 1687-1700
- Cordon-Obras C., Cano J., Gonzalez-Pacanowska D., Benito A., Navarro M., Bart J.-M. (2013) *Trypanosoma brucei gambiense* adaptation to different mammalian sera is associated with VSG expression site plasticity. *PLoS ONE* **8**
- Costes S. V., Chiolo I., Pluth J. M., Barcellos-Hoff M. H., Jakob B. (2010) Spatiotemporal characterization of ionizing radiation induced DNA damage

- foci and their relation to chromatin organization. *Mutation Research/Reviews in Mutation Research* **704**, 78-87
- Cross G. A. M., Kim H.-S., Wickstead B. (2014) Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. *Molecular and Biochemical Parasitology* **195**, 59-73
- Cross M., Taylor M. C., Borst P. (1998) Frequent loss of the active site during variant surface glycoprotein expression site switching *in vitro* in *Trypanosoma brucei*. *Molecular and Cellular Biology* **18**, 198-205
- Croteau D. L., Popuri V., Opresko P. L., Bohr V. A. (2014) Human RecQ helicases in DNA repair, recombination, and replication. *Annual Review of Biochemistry* **83**, 519-552
- Das S., Nozawa M., Klein J., Nei M. (2008) Evolutionary dynamics of the immunoglobulin heavy chain variable region genes in vertebrates. *Immunogenetics* **60**, 47-55
- Da Ines O., Degroote F., Goubely C., Amiard S., Gallego M. E., White C. I. (2013) Meiotic recombination in *Arabidopsis* is catalysed by DMC1, with RAD51 playing a supporting role. *PLoS Genetics* **9**
- Dalgaard J. Z., Klar A. J. (2000) Swi1 and Swi3 perform imprinting, pausing, and termination of DNA replication in *S. pombe*. *Cell* **102**, 745-751
- DeFazio L. G., Stansel R. M., Griffith J. D., Chu G. (2002) Synapsis of DNA ends by DNA-dependent protein kinase. *EMBO Journal* **21**, 3192-3200
- Deitsch K. W., Lukehart S. A., Stringer J. R. (2009) Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. *Nature Reviews Microbiology* **7**, 493-503
- Delange T., Kooter J. M., Luirink J., Borst P. (1985) Transcription of a transposed trypanosome surface-antigen gene starts upstream of the transposed segment. *EMBO Journal* **4**, 3299-3306
- Delange T., Kooter J. M., Michels P. A. M., Borst P. (1983) Telomere conversion in trypanosomes. *Nucleic Acids Research* **11**, 8149-8165
- Dherin C., Gueneau E., Francin M., Nunez M., Miron S., Liberti S. E., Rasmussen L. J., Zinn-Justin S., Gilquin B., Charbonnier J.B., Boiteux S. (2009) Characterization of a highly conserved binding site of Mlh1 required for exonuclease I-dependent mismatch repair. *Molecular and Cellular Biology* **29**, 907-918
- Dobson R., Stockdale C., Lapsley C., Wilkes J., McCulloch R. (2011) Interactions among *Trypanosoma brucei* RAD51 paralogues in DNA repair and antigenic variation. *Molecular Microbiology* **81**, 434-456
- Doe C. L., Ahn J. S., Dixon J., Whitby M. C. (2002) Mus81-Eme1 and Rqh1 involvement in processing stalled and collapsed replication forks. *Journal of Biological Chemistry* **277**, 32753-32759

- Doe C. L., Osman F., Dixon J., Whitby M. C. (2004) DNA repair by a Rad22-Mus81-dependent pathway that is independent of Rhp51. *Nucleic Acids Research* **32**, 5570-5581
- Doe C. L., Whitby M. C. (2004) The involvement of Srs2 in post-replication repair and homologous recombination in fission yeast. *Nucleic Acids Research* **32**, 1480-1491
- Downs J. A., Lowndes N. F., Jackson S. P. (2000) A role for *Saccharomyces cerevisiae* histone H2A in DNA repair. *Nature* **408**, 1001-1004
- Driessens N., Versteijhe S., Ghaddhab C., Burniat A., De Deken X., Van Sande J., Dumont J. E., Miot F., Corvilain B. (2009) Hydrogen peroxide induces DNA single- and double-strand breaks in thyroid cells and is therefore a potential mutagen for this organ. *Endocrine-related Cancer* **16**, 845-856
- Durand-Dubief M., Bastin P. (2003) TbAGO1, an argonaute protein required for RNA interference, is involved in mitosis and chromosome segregation in *Trypanosoma brucei*. *BMC Biology* **1**
- Dynan W. S., Yoo S. (1998) Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids. *Nucleic Acids Research* **26**, 1551-1559
- Egly J. M., Coin F. (2011) A history of TFIIH: two decades of molecular biology on a pivotal transcription/repair factor. *DNA Repair* **10**, 714-721
- El-Sayed N. M., Hegde P., Quackenbush J., Melville S. E., Donelson J. E. (2000) The African trypanosome genome. *International Journal for Parasitology* **30**, 329-345
- Esposito M. S. (1978) Evidence that spontaneous mitotic recombination occurs at 2-strand stage. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 4436-4440
- Falaschi A., Kornberg A. (1964) Antimetabolites affecting protein or nucleic acid synthesis - phleomycin inhibitor of DNA polymerase. *Federation Proceedings* **23**, 940-944
- Ferguson M. A. J., Homans S. W., Dwek R. A., Rademacher T. W. (1988) The glycosylphosphatidylinositol membrane anchor of the *Trypanosoma brucei* variant surface glycoprotein. *Biochemical Society Transactions* **16**, 265-268
- Figueiredo L. M., Cross G. A. M. (2010) Nucleosomes are depleted at the VSG expression site transcribed by RNA polymerase I in African trypanosomes. *Eukaryotic Cell* **9**, 148-154
- Figueiredo L. M., Janzen C. J., Cross G. A. M. (2008) A histone methyltransferase modulates antigenic variation in African trypanosomes. *PLoS Biology* **6**



- Fiorentini P., Huang K. N., Tishkoff D. X., Kolodner R. D., Symington L. S. (1997) Exonuclease I of *Saccharomyces cerevisiae* functions in mitotic recombination *in vivo* and *in vitro*. *Molecular and Cellular Biology* **17**, 2764-2773
- Fire A., Xu S. Q., Montgomery M. K., Kostas S. A., Driver S. E., Mello C. C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806-811
- Foury F., Kolodnyski J. (1983) Pif mutation blocks recombination between mitochondrial *rho*<sup>+</sup> and *rho*<sup>-</sup> genomes having tandemly arrayed repeat units in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* **80**, 5345-5349
- Foury F., Vandyck E. (1985) A Pif-dependent recombinogenic signal in the mitochondrial DNA of yeast. *EMBO Journal* **4**, 3525-3530
- Fu Y., Xiao W. (2003) Functional domains required for the *Saccharomyces cerevisiae* Mus81-Mms4 endonuclease complex formation and nuclear localization. *DNA Repair* **2**, 1435-1447
- Fuller-Pace F. V. (2006) DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. *Nucleic Acids Research* **34**, 4206-4215
- Gaillard P. H. L., Noguchi E., Shanahan P., Russell P. (2003) The endogenous Mus81-Eme1 complex resolves Holliday junctions by a nick and counternick mechanism. *Molecular Cell* **12**, 747-759
- Gao H., Chen X. B., McGowan C. H. (2003) Mus81 endonuclease localizes to nucleoli and to regions of DNA damage in human S-phase cells. *Molecular Biology of the Cell* **14**, 4826-4834
- García-Gómez S., Reyes A., Martínez-Jiménez María I., Chocrón E S., Mourón S., Terrados G., Powell C., Salido E., Méndez J., Holt Ian J., Blanco L. (2013) PrimPol, an archaic primase/polymerase operating in human cells. *Molecular Cell* **52**, 541-553
- Ginger M. L., Blundell P. A., Lewis A. M., Browitt A., Gunzl A., Barry J. D. (2002) *Ex vivo* and *in vitro* identification of a consensus promoter for VSG genes expressed by metacyclic-stage trypanosomes in the tsetse fly. *Eukaryotic Cell* **1**, 1000-1009
- Glover L., Alsford S., Beattie C., Horn D. (2007) Deletion of a trypanosome telomere leads to loss of silencing and progressive loss of terminal DNA in the absence of cell cycle arrest. *Nucleic Acids Research* **35**, 872-880
- Glover L., Alsford S., Horn D. (2013a) DNA break site at fragile subtelomeres determines probability and mechanism of antigenic variation in African trypanosomes. *PLoS Pathogens* **9**

- Glover L., Horn D. (2009) Site-specific DNA double-strand breaks greatly increase stable transformation efficiency in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **166**, 194-197
- Glover L., Horn D. (2012) Trypanosomal histone gamma H2A and the DNA damage response. *Molecular and Biochemical Parasitology* **183**, 78-83
- Glover L., Horn D. (2014) Locus-specific control of DNA resection and suppression of subtelomeric VSG recombination by HAT3 in the African trypanosome. *Nucleic Acids Research* **42**, 12600-12613
- Glover L., Hutchinson S., Alsford S., McCulloch R., Field M. C., Horn D. (2013b) Antigenic variation in African trypanosomes: the importance of chromosomal and nuclear context in VSG expression control. *Cellular Microbiology* **15**, 1984-1993
- Glover L., Jun J., Horn D. (2011) Microhomology-mediated deletion and gene conversion in African trypanosomes. *Nucleic Acids Research* **39**, 1372-1380
- Glover L., McCulloch R., Horn D. (2008) Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. *Nucleic Acids Research* **36**, 2608-2618
- Gorbalenya A. E., Koonin E. V. (1993) Helicases - amino-acid-sequence comparisons and structure-function-relationships. *Current Opinion in Structural Biology* **3**, 419-429
- Gravel S., Chapman J. R., Magill C., Jackson S. P. (2008) DNA helicases Sgs1 and BLM promote DNA double-strand break resection. *Genes & Development* **22**, 2767-2772
- Grawunder U., Wilm M., Wu X. T., Kulesza P., Wilson T. E., Mann M., Lieber M. R. (1997) Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* **388**, 492-495
- Greinert R., Volkmer B., Henning S., Breitbart E. W., Greulich K. O., Cardoso M. C., Rapp A. (2012) UVA-induced DNA double-strand breaks result from the repair of clustered oxidative DNA damages. *Nucleic Acids Research* **40**, 10263-10273
- Groth P., Auslander S., Majumder M. M., Schultz N., Johansson F., Petermann E., Helleday T. (2010) Methylated DNA causes a physical block to replication forks independently of damage signalling, O(6)-methylguanine or DNA single-strand breaks and results in DNA damage. *Journal of Molecular Biology* **402**, 70-82
- Guizetti J., Scherf A. (2013) Silence, activate, poise and switch! Mechanisms of antigenic variation in *Plasmodium falciparum*. *Cellular Microbiology* **15**, 718-726
- Gunzl A., Bruderer T., Laufer G., Schimanski B., Tu L. C., Chung H. M., Lee P. T., Lee M. G. S. (2003) RNA polymerase I transcribes procyclin genes and

variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Eukaryotic Cell* **2**, 542-551

Günzl A., Kirkham J. K., Nguyen T. N., Badjatia N., Park S. H. (2015) Mono-allelic VSG expression by RNA polymerase I in *Trypanosoma brucei*: Expression site control from both ends? *Gene* **556**, 68-73

Guo R.-B., Rigolet P., Ren H., Zhang B., Zhang X.-D., Dou S.-X., Wang P.-Y., Amor-Gueret M., Xi X. G. (2007) Structural and functional analyses of disease-causing missense mutations in Bloom syndrome protein. *Nucleic Acids Research* **35**, 6297-6310

Haaf T., Golub E. I., Reddy G., Radding C. M., Ward D. C. (1995) Nuclear foci of mammalian RAD51 recombination protein in somatic-cells after DNA-damage and its localization in synaptonemal complexes. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 2298-2302

Haber J. E. (1998) Mating-type gene switching in *Saccharomyces cerevisiae*. *Annu Rev Genet* **32**, 561-599

Hajduk S. L., Hager K., Esko J. D. (1992) High density lipoprotein-mediated lysis of trypanosomes. *Parasitology Today* **8**, 95-98

Hall J. P. J. (2012) Mosaic VSGs in *Trypanosoma brucei* antigenic variation. PhD Thesis, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow,

Hall J. P. J., Wang H., Barry J. D. (2013) Mosaic VSGs and the scale of *Trypanosoma brucei* antigenic variation. *PLoS Pathogens* **9**

Hanada K., Budzowska M., Davies S. L., van Drunen E., Onizawa H., Beverloo H. B., Maas A., Essers J., Hickson I. D., Kanaar R. (2007) The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks. *Nature Structural & Molecular Biology* **14**, 1096-1104

Hanada K., Budzowska M., Modesti M., Maas A., Wyman C., Essers J., Kanaar R. (2006) The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinks into double-strands breaks. *EMBO Journal* **25**, 4921-4932

Harrington J. M., Howell S., Hajduk S. L. (2009) Membrane permeabilization by trypanosome lytic factor, a cytolytic human high density lipoprotein. *Journal of Biological Chemistry* **284**, 13505-13512

Hartley C. L. (2008) Characterisation of factors that regulate homologous recombination and antigenic variation in *Trypanosoma brucei*. PhD Thesis, University of Glasgow

Hartley C. L., McCulloch R. (2008) *Trypanosoma brucei* BRCA2 acts in antigenic variation and has undergone a recent expansion in BRC repeat number

that is important during homologous recombination. *Molecular Microbiology* **68**, 1237-1251

- Hartung F., Suer S., Knoll A., Wurz-Wildersinn R., Puchta H. (2008) Topoisomerase 3alpha and RMI1 suppress somatic crossovers and are essential for resolution of meiotic recombination intermediates in *Arabidopsis thaliana*. *PLoS Genetics* **4**
- Helmrich A., Ballarino M., Tora L. (2011) Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Molecular Cell* **44**, 966-977
- Hertz-Fowler C., Figueiredo L. M., Quail M. A., Becker M., Jackson A., Bason N., Brooks K., Churcher C., Fahrenberg S., Goodhead I., Heath P., Kartvelishvili M., Mungall K., Harris D., Hauser H., Sanders M., Saunders D., Seeger K., Sharp S., Taylor J. E., Walker D., White B., Young R., Cross G. A. M., Rudenko G., Barry J. D., Louis E. J., Berriman M. (2008) Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *PLoS ONE* **3**
- Heyman R. A., Borrelli E., Lesley J., Anderson D., Richman D. D., Baird S. M., Hyman R., Evans R. M. (1989) Thymidine kinase obliteration - creation of transgenic mice with controlled immune-deficiency. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 2698-2702
- Hickson I. D. (2003) RecQ helicases: caretakers of the genome. *Nature Reviews Cancer* **3**, 169-178
- Hickson I. D., Mankouri H. W. (2011) Processing of homologous recombination repair intermediates by the Sgs1-Top3-Rmi1 and Mus81-Mms4 complexes. *Cell Cycle* **10**, 3078-3085
- Higgins M. K., Tkachenko O., Brown A., Reed J., Raper J., Carrington M. (2013) Structure of the trypanosome haptoglobin-hemoglobin receptor and implications for nutrient uptake and innate immunity. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 1905-1910
- Hill S. A., Davies J. K. (2009) Pilin gene variation in *Neisseria gonorrhoeae*: reassessing the old paradigms. *FEMS Microbiology Reviews* **33**, 521-530
- Hirumi H., Hirumi K. (1989) Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum-protein without feeder cell-layers. *Journal of Parasitology* **75**, 985-989
- Holliday R. (1964) Mechanism for gene conversion in fungi. *Genetical Research* **5**, 285-307
- Hollingsworth N. M., Brill S. J. (2004) The Mus81 solution to resolution: generating meiotic crossovers without Holliday junctions. *Genes & Development* **18**, 117-125

- Holloway J. K., Booth J., Edelman W., McGowan C. H., Cohen P. E. (2008) MUS81 generates a subset of MLH1-MLH3-independent crossovers in mammalian meiosis. *PLoS Genetics* 4
- Horn D. (2014) Antigenic variation in African trypanosomes. *Molecular and Biochemical Parasitology* 195, 123-129
- Horn D., Barry J. D. (2005) The central roles of telomeres and subtelomeres in antigenic variation in African trypanosomes. *Chromosome Research* 13, 525-533
- Hsiang Y. H., Hertzberg R., Hecht S., Liu L. F. (1985) Camptothecin induces protein-linked DNA breaks via mammalian dna topoisomerase- $\alpha$ . *Journal of Biological Chemistry* 260, 4873-4878
- Hunter S., Jones P., Mitchell A., Apweiler R., Attwood T. K., Bateman A., Bernard T., Binns D., Bork P., Burge S., de Castro E., Coggill P., Corbett M., Das U., Daugherty L., Duquenne L., Finn R. D., Fraser M., Gough J., Haft D., Hulo N., Kahn D., Kelly E., Letunic I., Lonsdale D., Lopez R., Madera M., Maslen J., McAnulla C., McDowall J., McMenamin C., Mi H., Mutowo-Muellenet P., Mulder N., Natale D., Orengo C., Pesseat S., Punta M., Quinn A. F., Rivoire C., Sangrador-Vegas A., Selengut J. D., Sigrist C. J. A., Scheremetjew M., Tate J., Thimmajanarathan M., Thomas P. D., Wu C. H., Yeats C., Yong S.-Y. (2012) InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Research* 40, 306-312
- Huppert J. L. (2008) Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chemical Society Reviews* 37, 1375-1384
- Hutchinson O. C., Smith W., Jones N. G., Chattopadhyay A., Welburn S. C., Carrington M. (2003) VSG structure: similar N-terminal domains can form functional VSGs with different types of C-terminal domain. *Molecular and Biochemical Parasitology* 130, 127-131
- Hyun M., Park S., Kim E., Kim D.-H., Lee S.-J., Koo H.-S., Seo Y.-S., Ahn B. (2012) Physical and functional interactions of *Caenorhabditis elegans* WRN-1 helicase with RPA-1. *Biochemistry* 51, 1336-1345
- Im J. S., Ki S. H., Farina A., Jung S., Hurwitz J., Lee J. K. (2009) Assembly of the Cdc45-Mcm2-7-GINS complex in human cells requires the Ctf4/And-1, RecQL4, and Mcm10 proteins. *Proceedings of the National Academy of Sciences of the United States of America* 106, 15628-15632
- Imamura O., Fujita K., Shimamoto A., Tanabe H., Takeda S., Furuichi Y., Matsumoto T. (2001) Bloom helicase is involved in DNA surveillance in early S phase in vertebrate cells. *Oncogene* 20, 1143-1151
- Imboden M. A., Laird P. W., Affolter M., Seebeck T. (1987) Transcription of the intergenic regions of the tubulin gene-cluster of *Trypanosoma brucei* - evidence for a polycistronic transcription unit in a eukaryote. *Nucleic Acids Research* 15, 7357-7368

- Interthal H., Heyer W. D. (2000)** MUS81 encodes a novel Helix-hairpin-Helix protein involved in the response to UV- and methylation-induced DNA damage in *Saccharomyces cerevisiae*. *Molecular and General Genetics* **263**, 812-827
- Ira G., Haber J. E. (2002)** Characterization of RAD51-independent break-induced replication that acts preferentially with short homologous sequences. *Molecular and Cellular Biology* **22**, 6384-6392
- Ira G., Malkova A., Liberi G., Foiani M., Haber J. E. (2003)** Srs2 and Sgs1-Top3 suppress crossovers during double-strand break repair in yeast. *Cell* **115**, 401-411
- Ivanov E. L., Sugawara N., FishmanLobell J., Haber J. E. (1996)** Genetic requirements for the single-strand annealing pathway of double-strand break repair in *Saccharomyces cerevisiae*. *Genetics* **142**, 693-704
- Ivessa A. S., Zhou J. Q., Schulz V. P., Monson E. K., Zakian V. A. (2002)** *Saccharomyces* Rrm3p, a 5' to 3' DNA helicase that promotes replication fork progression through telomeric and subtelomeric DNA. *Genes & Development* **16**, 1383-1396
- Jackson S. P. (2002)** Sensing and repairing DNA double-strand breaks. *Carcinogenesis* **23**, 687-696
- Jakupciak J. P., Wells R. D. (1999)** Genetic instabilities in (CTG center dot CAG) repeats occur by recombination. *Journal of Biological Chemistry* **274**, 23468-23479
- Jamonneau V., Ilboudo H., Kabore J., Kaba D., Koffi M., Solano P., Garcia A., Courtin D., Laveissiere C., Lingue K., Buscher P., Bucheton B. (2012)** Untreated human infections by *Trypanosoma brucei gambiense* are not 100% fatal. *PLoS Neglected Tropical Diseases* **6**
- Jehi S. E., Li X., Sandhu R., Ye F., Benmerzouga I., Zhang M., Zhao Y., Li B. (2014a)** Suppression of subtelomeric VSG switching by *Trypanosoma brucei* TRF requires its TTAGGG repeat-binding activity. *Nucleic Acids Research* **42**, 12899-12911
- Jehi S. E., Wu F., Li B. (2014b)** *Trypanosoma brucei* TIF2 suppresses VSG switching by maintaining subtelomere integrity. *Cell Res* **24**, 870-885
- Jensen R. B., Carreira A., Kowalczykowski S. C. (2010)** Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**, 678-683
- Jeong Y. S., Kang Y. L., Lim K. H., Lee M. H., Lee J., Koo H. S. (2003)** Deficiency of *Caenorhabditis elegans* RecQ5 homologue reduces life span and increases sensitivity to ionizing radiation. *DNA Repair* **2**, 1309-1319
- Johnson F. B., Lombard D. B., Neff N. F., Mastrangelo M. A., Dewolf W., Ellis N. A., Marciniak R. A., Yin Y. H., Jaenisch R., Guarente L. (2000)** Association of the bloom syndrome protein with topoisomerase III alpha in somatic and meiotic cells. *Cancer Research* **60**, 1162-1167

- Jones N. G. (2013) Validating protein kinases of *Trypanosoma brucei* as drug targets. Doctor of Philosophy Thesis, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow,
- Jones N. G., Thomas E. B., Brown E., Dickens N. J., Hammarton T. C., Mottram J. C. (2014) Regulators of *Trypanosoma brucei* cell cycle progression and differentiation identified using a kinome-wide RNAi screen. *PLoS Pathogens* 10
- Jung H., Lee J. A., Choi S., Lee H., Ahn B. (2014) Characterization of the *Caenorhabditis elegans* HIM-6/BLM helicase: unwinding recombination intermediates. *PLoS ONE* 9
- Kai M., Boddy M. N., Russell P., Wang T. S. F. (2005) Replication checkpoint kinase Cds1 regulates Mus81 to reserve genome integrity during replication stress. *Genes & Development* 19, 919-932
- Kaliraman V., Mullen J. R., Fricke W. M., Bastin-Shanower S. A., Brill S. J. (2001) Functional overlap between Sgs1-Top3 and the Mms4-Mus81 endonuclease. *Genes & Development* 15, 2730-2740
- Kamileri I., Karakasilioti I., Garinis G. A. (2012) Nucleotide excision repair: new tricks with old bricks. *Trends in Genetics* 28, 566-573
- Kamper S. M., Barbet A. F. (1992) Surface epitope variation via mosaic gene formation is potential key to long-term survival of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 53, 33-44
- Kang S. M., Ohshima K., Shimizu M., Amirhaeri S., Wells R. D. (1995) Pausing of DNA-synthesis *in vitro* at specific loci in CTG and CGG triplet repeats from human hereditary-disease genes. *Journal of Biological Chemistry* 270, 27014-27021
- Karmakar P., Seki M., Kanamori M., Hashiguchi K., Ohtsuki M., Murata E., Inoue E., Tada S., Lan L., Yasui A., Enomoto T. (2006) BLM is an early responder to DNA double-strand breaks. *Biochemical and Biophysical Research Communications* 348, 62-69
- Kaykov A., Holmes A. M., Arcangioli B. (2004) Formation, maintenance and consequences of the imprint at the mating-type locus in fission yeast. *EMBO Journal* 23, 930-938
- Keogh M. C., Kim J. A., Downey M., Fillingham J., Chowdhury D., Harrison J. C., Onishi M., Datta N., Galicia S., Emili A., Lieberman J., Shen X. T., Buratowski S., Haber J. E., Durocher D., Greenblatt J. F., Krogan N. J. (2006) A phosphatase complex that dephosphorylates gamma H2AX regulates DNA damage checkpoint recovery. *Nature* 439, 497-501
- Khanna K. K., Jackson S. P. (2001) DNA double-strand breaks: signaling, repair and the cancer connection. *Nature Genetics* 27, 247-254

- Kikuchi K., Narita T., Pham V. T., Iijima J., Hirota K., Keka I. S., Okawa K., Hori T., Fukagawa T., Essers J., Kanaar R., Whitby M. C., Sugasawa K., Taniguchi Y., Kitagawa K., Takeda S. (2013) Structure-specific endonucleases Xpf and Mus81 play overlapping but essential roles in DNA repair by homologous recombination. *Cancer Research* **73**, 4362-4371
- Kim H.S., Cross G. A. M. (2010) TOPO3alpha influences antigenic variation by monitoring expression-site-associated VSG switching in *Trypanosoma brucei*. *PLoS pathogens* **6**, e1000992
- Kim H.S., Cross G. A. M. (2011) Identification of *Trypanosoma brucei* RMI1/BLAP75 homologue and its roles in antigenic variation. *PLoS ONE* **6**,
- Kitano K., Kim S. Y., Hakoshima T. (2010) Structural basis for DNA strand separation by the unconventional winged-helix domain of RecQ helicase WRN. *Structure* **18**, 177-187
- Klar A. J. S. (2007) Lessons learned from studies of fission yeast mating-type switching and silencing. In *Annual Review of Genetics* Vol. 41, pp 213-236.
- Klingbeil M. M., Motyka S. A., Englund P. T. (2002) Multiple mitochondrial DNA polymerases in *Trypanosoma brucei*. *Molecular Cell* **10**, 175-186
- Kolev N. G., Ramey-Butler K., Cross G. A. M., Ullu E., Tschudi C. (2012) Developmental progression to infectivity in *Trypanosoma brucei* triggered by an RNA-binding protein. *Science* **338**, 1352-1353
- Kolev N. G., Tschudi C., Ullu E. (2011) RNA interference in protozoan parasites: achievements and challenges. *Eukaryotic Cell* **10**, 1156-1163
- Krejci L., Van Komen S., Li Y., Villemain J., Reddy M. S., Klein H., Ellenberger T., Sung P. (2003) DNA helicase Srs2 disrupts the Rad51 presynaptic filament. *Nature* **423**, 305-309
- Kuper J., Braun C., Elias A., Michels G., Sauer F., Schmitt D. R., Poterszman A., Egly J.-M., Kisker C. (2014) In TFIIH, XPD helicase is exclusively devoted to DNA repair. *PLoS Biology* **12**
- Kurosawa A., Saito S., So S., Hashimoto M., Iwabuchi K., Watabe H., Adachi N. (2013) DNA Ligase IV and Artemis act cooperatively to suppress homologous recombination in human cells: implications for DNA double-strand break repair. *PLoS ONE* **8**,
- Kusumoto R., Dawut L., Marchetti C., Lee J. W., Vindigni A., Ramsden D., Bohr V. A. (2008) Werner protein cooperates with the XRCC4-DNA ligase IV complex in end-processing. *Biochemistry* **47**, 7548-7556
- La Greca F., Magez S. (2011) Vaccination against trypanosomiasis Can it be done or is the trypanosome truly the ultimate immune destroyer and escape artist? *Human Vaccines* **7**, 1225-1233
- Labib K., Hodgson B. (2007) Replication fork barriers: pausing for a break or stalling for time? *EMBO Reports* **8**, 346-353



- Labib K., Tercero J. A., Diffley J. F. X. (2000) Uninterrupted MCM2-7 function required for DNA replication fork progression. *Science* **288**, 1643-1647
- Lahaye A., Stahl H., Thines-Sempoux D., Foury F. (1991) PIF1: a DNA helicase in yeast mitochondria. *EMBO Journal* **10**, 997-1007
- Lamont G. S., Tucker R. S., Cross G. A. (1986) Analysis of antigen switching rates in *Trypanosoma brucei*. *Parasitology* **92** ( Part 2), 355-367
- Lan L., Nakajima S., Komatsu K., Nussenzweig A., Shimamoto A., Oshima J., Yasui A. (2005) Accumulation of Werner protein at DNA double-strand breaks in human cells. *Journal of Cell Science* **118**, 4153-4162
- Landeira D., Bart J. M., Van Tyne D., Navarro M. (2009) Cohesin regulates VSG monoallelic expression in trypanosomes. *Journal of Cell Biology* **186**, 243-254
- Langland G., Elliott J., Li Y. L., Creaney J., Dixon K., Groden J. (2002) The BLM helicase is necessary for normal DNA double-strand break repair. *Cancer Research* **62**, 2766-2770
- Langland G., Kordich J., Creaney J., Goss K. H., Lillard-Wetherell K., Bebenek K., Kunkel T. A., Groden J. (2001) The Bloom's syndrome protein (BLM) interacts with MLH1 but is not required for DNA mismatch repair. *Journal of Biological Chemistry* **276**, 30031-30035
- Lebel M., Spillare E. A., Harris C. C., Leder P. (1999) The Werner syndrome gene product co-purifies with the DNA replication complex and interacts with PCNA and topoisomerase I. *Journal of Biological Chemistry* **274**, 37795-37799
- Lee M. G., Van der Ploeg L. H. T. (1987) Frequent independent duplicative transpositions activate a single VSG gene. *Molecular and Cellular Biology* **7**, 357-364
- Lees-Miller S. P., Meek K. (2003a) Repair of DNA double strand breaks by non-homologous end joining. *Biochimie* **85**, 1161-1173
- Lees-Miller S. P., Meek K. (2003b) Repair of DNA double strand breaks by non-homologous end joining. *Biochimie* **85**, 1161-1173
- Li J.-R., Yu T.-Y., Chien I. C., Lu C.-Y., Lin J.-J., Li H.-W. (2014) Pif1 regulates telomere length by preferentially removing telomerase from long telomere ends. *Nucleic Acids Research* **42**, 8527-8536
- Linder P. (2006) Dead-box proteins: a family affair—active and passive players in RNP-remodeling. *Nucleic Acids Research* **34**, 4168-4180
- Liu A. Y. C., Vanderploeg L. H. T., Rijsewijk F. A. M., Borst P. (1983) The transposition unit of variant surface glycoprotein gene-118 of *Trypanosoma brucei* - presence of repeated elements at its border and absence of promoter-associated sequences. *Journal of Molecular Biology* **167**, 57-75

- Liu B., Wang J., Yaffe N., Lindsay M. E., Zhao Z., Zick A., Shlomai J., Englund P. T. (2009a) Trypanosomes have six mitochondrial DNA helicases with one controlling kinetoplast maxicircle replication. *Molecular Cell* **35**, 490-501
- Liu B., Wang J., Yildirim G., Englund P. T. (2009b) TbPIF5 is a *Trypanosoma brucei* mitochondrial DNA helicase involved in processing of minicircle Okazaki fragments. *PLoS Pathogens* **5**
- Liu B., Yildirim G., Wang J., Tolun G., Griffith J. D., Englund P. T. (2010a) TbPIF1, a *Trypanosoma brucei* mitochondrial DNA helicase, is essential for kinetoplast minicircle replication. *Journal of Biological Chemistry* **285**, 7056-7066
- Liu H., Yan P., Fanning E. (2015) Human DNA helicase B functions in cellular homologous recombination and stimulates Rad51-mediated 5'-3' heteroduplex extension *in vitro*. *PLoS ONE* **10**
- Liu J., Doty T., Gibson B., Heyer W.-D. (2010b) Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nature Structural & Molecular Biology* **17**, 1260-1262
- Liu L. F., Desai S. D., Li T. K., Mao Y., Sun M., Sim S. P. (2000) Mechanism of action of camptothecin. *Annals of the New York Academy of Sciences* **922**, 1-10
- Llorente B., Smith C. E., Symington L. S. (2008) Break-induced replication: What is it and what is it for? *Cell Cycle* **7**, 859-864
- Lonn U., Lonn S., Nylen U., Winblad G., German J. (1990) An abnormal profile of DNA-replication intermediates in Blooms syndrome. *Cancer Research* **50**, 3141-3145
- Lopes J., Piazza A., Bermejo R., Kriegsman B., Colosio A., Teulade-Fichou M.-P., Foiani M., Nicolas A. (2011) G-quadruplex-induced instability during leading-strand replication. *EMBO Journal* **30**, 4033-4046
- Lopez-Farfan D., Bart J. M., Rojas-Barros D. I., Navarro M. (2014) SUMOylation by the E3 ligase TbSIZ1/PIAS1 positively regulates VSG expression in *Trypanosoma brucei*. *PLoS Pathogens* **10**, 18
- Louis E. J., Vershinin A. V. (2005) Chromosome ends: different sequences may provide conserved functions. *Bioessays* **27**, 685-697
- Lukes J., Guilbride D. L., Votypka J., Zikova A., Benne R., Englund P. T. (2002) Kinetoplast DNA network: Evolution of an improbable structure. *Eukaryotic Cell* **1**, 495-502
- Lundin C., North M., Erixon K., Walters K., Jenssen D., Goldman A. S. H., Helleday T. (2005) Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable *in vivo* DNA double-strand breaks. *Nucleic Acids Research* **33**, 3799-3811

- Ma W., Westmoreland J. W., Gordenin D. A., Resnick M. A. (2011) Alkylation base damage is converted into repairable double-strand breaks and complex intermediates in G2 cells lacking AP endonuclease. *PLoS Genetics* **7**
- MacGregor P., Savill Nicholas J., Hall D., Matthews Keith R. (2011) Transmission stages dominate trypanosome within-host dynamics during chronic infections. *Cell Host & Microbe* **9**, 310-318
- MacGregor P., Szoer B., Savill N. J., Matthews K. R. (2012) Trypanosomal immune evasion, chronicity and transmission: an elegant balancing act. *Nature Reviews Microbiology* **10**, 431-438
- Machado C. R., Vieira-da-Rocha J. P., Mendes I. C., Rajao M. A., Marcello L., Bitar M., Drummond M. G., Grynberg P., Oliveira D. A. A., Marques C., Van Houten B., McCulloch R. (2014) Nucleotide excision repair in *Trypanosoma brucei*: specialization of transcription-coupled repair due to multigenic transcription. *Molecular Microbiology* **92**, 756-776
- Makovets S., Blackburn E. H. (2009) DNA damage signalling prevents deleterious telomere addition at DNA breaks. *Nature Cell Biology* **11**, 1383-1386
- Malkova A., Ivanov E. L., Haber J. E. (1996) Double-strand break repair in the absence of RAD51 in yeast: A possible role for break-induced DNA replication. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 7131-7136
- Malkova A., Klein F., Leung W. Y., Haber J. E. (2000) HO endonuclease-induced recombination in yeast meiosis resembles Spo11-induced events. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 14500-14505
- Mansour S. L., Thomas K. R., Capecchi M. R. (1988) Disruption of the proto-oncogene int-2 in mouse embryo-derived stem-cells - a general strategy for targeting mutations to non-selectable genes. *Nature* **336**, 348-352
- Mao F. J., Sidorova J. M., Lauper J. M., Emond M. J., Monnat R. J. (2010) The human WRN and BLM RecQ helicases differentially regulate cell proliferation and survival after chemotherapeutic DNA damage. *Cancer Research* **70**, 6548-6555
- Marcello L., Barry J. D. (2007) Analysis of the VSG gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. *Genome Research* **17**, 1344-1352
- Marques C. A. (2015) Initiation of nuclear DNA replication in *Trypanosoma brucei* and *Leishmania*. PhD Thesis, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow

- Matsuda K., Makise M., Sueyasu Y., Takehara M., Asano T., Mizushima T. (2007)** Yeast two-hybrid analysis of the origin recognition complex of *Saccharomyces cerevisiae*: interaction between subunits and identification of binding proteins. *FEMS Yeast Research* **7**, 1263-1269
- Matthews K. R., Gull K. (1997)** Commitment to differentiation and cell cycle re-entry are coincident but separable events in the transformation of African trypanosomes from their bloodstream to their insect form. *Journal of Cell Science* **110**, 2609-2618
- Matthews K. R., Shiels P. G., Graham S. V., Cowan C., Barry J. D. (1990)** Duplicative activation mechanisms of 2 trypanosome telomeric VSG genes with structurally simple 5' flanks. *Nucleic Acids Research* **18**, 7219-7227
- McCulloch R., Barry J. D. (1999)** A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation. *Genes & Development* **13**, 2875-2888
- McCulloch R., Horn D. (2009)** What has DNA sequencing revealed about the VSG expression sites of African trypanosomes? *Trends in Parasitology* **25**, 359-363
- McCulloch R., Rudenko G., Borst P. (1997)** Gene conversions mediating antigenic variation in *Trypanosoma brucei* can occur in variant surface glycoprotein expression sites lacking 70-base-pair repeat sequences. *Molecular and Cellular Biology* **17**, 833-843
- McEachern M. J., Haber J. E. (2006)** Break-induced replication and recombinational telomere elongation in yeast. *Annual Review of Biochemistry* **75**, 111-135.
- McElhinny S. A. N., Snowden C. M., McCarville J., Ramsden D. A. (2000)** Ku recruits the XRCC4-ligase IV complex to DNA ends. *Molecular and Cellular Biology* **20**, 2996-3003
- McVey M., Lee S. E. (2008)** MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends in Genetics* **24**, 529-538
- Mefford H. C., Cook J., Gospe S. M., Jr. (2012)** Epilepsy due to 20q13.33 subtelomere deletion masquerading as pyridoxine-dependent epilepsy. *American Journal of Medical Genetics Part A* **158A**, 3190-3195
- Mefford H. C., Linardopoulou E., Coil D., van den Engh G., Trask B. J. (2001)** Comparative sequencing of a multicopy subtelomeric region containing olfactory receptor genes reveals multiple interactions between non-homologous chromosomes. *Human Molecular Genetics* **10**, 2363-2372
- Merrih H., Machon C., Grainger W. H., Grossman A. D., Soutanas P. (2011)** Co-directional replication-transcription conflicts lead to replication restart. *Nature* **470**, 554-557

- Milord F., Pépin J., Ethier L., Loko L., Mpia B. (1992) Efficacy and toxicity of eflornithine for treatment of *Trypanosoma brucei* gambiense sleeping sickness. *The Lancet* **340**, 652-655
- Mimitou E. P., Symington L. S. (2008) Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* **455**, 770-773
- Monici M. (2005) Cell and tissue autofluorescence research and diagnostic applications. *Biotechnology Annual Review* **11**, 227-256
- Morel F., Renoux M., Lachaurne P., Alziari S. (2008) Bleomycin-induced double-strand breaks in mitochondrial DNA of *Drosophila* cells are repaired. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* **637**, 111-117
- Morozov V., Mushegian A. R., Koonin E. V., Bork P. (1997) A putative nucleic acid-binding domain in Bloom's and Werner's syndrome helicases. *Trends in Biochemical Sciences* **22**, 417-418
- Morrison L., Majiwa P., Read A., Barry J. (2005) Probabilistic order in antigenic variation of *Trypanosoma brucei*. *International Journal for Parasitology* **35**, 961-972
- Morrison L. J., Marcello L., McCulloch R. (2009) Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. *Cellular Microbiology* **11**, 1724-1734
- Mullen J. R., Kaliraman V., Brill S. J. (2000) Bipartite structure of the SGS1 DNA helicase in *Saccharomyces cerevisiae*. *Genetics* **154**, 1101-1114
- Mullen J. R., Kaliraman V., Ibrahim S. S., Brill S. J. (2001) Requirement for three novel protein complexes in the absence of the Sgs1 DNA helicase in *Saccharomyces cerevisiae*. *Genetics* **157**, 103-118
- Mullen J. R., Nallaseth F. S., Lan Y. Q., Slagle C. E., Brill S. J. (2005) Yeast Rmi1/Nce4 controls genome stability as a subunit of the Sgs1-Top3 complex. *Molecular and Cellular Biology* **25**, 4476-4487
- Myung K., Datta A., Chen C., Kolodner R. D. (2001) SGS1, the *Saccharomyces cerevisiae* homologue of BLM and WRN, suppresses genome instability and homeologous recombination. *Nature Genetics* **27**, 113-116
- Nakada S., Chen G. I., Gingras A.-C., Durocher D. (2008) PP4 is a gamma H2AX phosphatase required for recovery from the DNA damage checkpoint. *EMBO Reports* **9**, 1019-1026
- Nakamura T. M., Du L. L., Redon C., Russell P. (2004) Histone H2A phosphorylation controls Crb2 recruitment at DNA breaks, maintains checkpoint arrest, and influences DNA repair in fission yeast. *Molecular and Cellular Biology* **24**, 6215-6230
- Nandakurnar J., Cech TR. (2013) Finding the end: recruitment of telomerase to telomeres. *Nature Reviews Molecular Cell Biology* **14**, 69-82

- Nassif N., Penney J., Pal S., Engels W. R., Gloor G. B. (1994) Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Molecular and Cellular Biology* **14**, 1613-1625
- Navarro M., Gull K. (2001) A Pol I transcriptional body associated with VSG mono-allelic expression in *Trypanosoma brucei*. *Nature* **414**, 759-763
- Ngo H., Tschudi C., Gull K., Ullu E. (1998) Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 14687-14692
- Nguyen T. N., Mueller L. S. M., Park S. H., Siegel T. N., Guenzl A. (2014) Promoter occupancy of the basal class I transcription factor A differs strongly between active and silent VSG expression sites in *Trypanosoma brucei*. *Nucleic Acids Research* **42**, 3164-3176
- Nicolette M. L., Lee K., Guo Z., Rani M., Chow J. M., Lee S. E., Paull T. T. (2010) Mre11-Rad50-Xrs2 and Sae2 promote 5' strand resection of DNA double-strand breaks. *Nature Structural & Molecular Biology* **17**, 1478-1485
- Nimonkar A. V., Genschel J., Kinoshita E., Polaczek P., Campbell J. L., Wyman C., Modrich P., Kowalczykowski S. C. (2011) BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes & Development* **25**, 350-362
- Nomura Y., Adachi N., Koyama H. (2007) Human Mus81 and FANCB independently contribute to repair of DNA damage during replication. *Genes to Cells* **12**, 1111-1122
- Nussenzweig A., Nussenzweig M. C. (2007) A backup DNA repair pathway moves to the forefront. *Cell* **131**, 223-225
- O'Rourke T. W., Doudican N. A., Mackereth M. D., Doetsch P. W., Shadel G. S. (2002) Mitochondrial dysfunction due to oxidative mitochondrial DNA damage is reduced through cooperative actions of diverse proteins. *Molecular and Cellular Biology* **22**, 4086-4093
- Ogrunc M., Sancar A. (2003) Identification and characterization of human MUS81-MMS4 structure-specific endonuclease. *Journal of Biological Chemistry* **278**, 21715-21720
- Ohshima K., Kang S., Larson J. E., Wells R. D. (1996a) Cloning, characterization, and properties of seven triplet repeat DNA sequences. *Journal of Biological Chemistry* **271**, 16773-16783
- Ohshima K., Kang S., Larson J. E., Wells R. D. (1996b) TTA center dot TAA triplet repeats in plasmids form a non-H bonded structure. *Journal of Biological Chemistry* **271**, 16784-16791

- Osman F., Dixon J., Doe C. L., Whitby M. C. (2003) Generating crossovers by resolution of nicked Holliday junctions: A role for Mus81-Eme1 in meiosis. *Molecular Cell* **12**, 761-774
- Paeschke K., Bochman M. L., Daniela Garcia P., Cejka P., Friedman K. L., Kowalczykowski S. C., Zakian V. A. (2013) Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature* **497**, 458-462
- Palmer G. H., Bankhead T., Lukehart S. A. (2009) 'Nothing is permanent but change' - antigenic variation in persistent bacterial pathogens. *Cellular Microbiology* **11**, 1697-1705
- Paques F., Haber J. E. (1999) Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiology and Molecular Biology Reviews* **63**, 349-404
- Pardo B., Aguilera A. (2012) Complex chromosomal rearrangements mediated by break-induced replication involve structure-selective endonucleases. *PLoS Genetics* **8**
- Parvathaneni S., Stortchevoi A., Sommers J. A., Brosh R. M., Sharma S. (2013) Human RECQ1 interacts with Ku70/80 and modulates DNA end-joining of double-strand breaks. *PLoS ONE* **8**
- Patrick K. L., Shi H., Kolev N. G., Ersfeld K., Tschudi C., Ullu E. (2009) Distinct and overlapping roles for two Dicer-like proteins in the RNA interference pathways of the ancient eukaryote *Trypanosoma brucei*. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 17933-17938
- Paul T. T., Gellert M. (1998) The 3' to 5' exonuclease activity of Mre11 facilitates repair of DNA double-strand breaks. *Molecular Cell* **1**, 969-979
- Pays E., Guyaux M., Aerts D., Vanmeirvenne N., Steinert M. (1985) Telomeric reciprocal recombination as a possible mechanism for antigenic variation in trypanosomes. *Nature* **316**, 562-564
- Pays E., Lips S., Nolan D., Vanhamme L., Perez-Morga D. (2001) The VSG expression sites of *Trypanosoma brucei*: multipurpose tools for the adaptation of the parasite to mammalian hosts. *Molecular and Biochemical Parasitology* **114**, 1-16
- Pays E., Van Assel S., Laurent M., Darville M., Vervoort T., VanMeirvenne N., Steinert M. (1983) Gene conversion as a mechanism for antigenic variation in trypanosomes. *Cell* **34**, 371-381
- Peacock L., Bailey M., Carrington M., Gibson W. (2014) Meiosis and haploid gametes in the pathogen *Trypanosoma brucei*. *Current Biology* **24**, 181-186
- Peacock L., Ferris V., Sharma R., Sunter J., Bailey M., Carrington M., Gibson W. (2011) Identification of the meiotic life cycle stage of *Trypanosoma*

*brucei* in the tsetse fly. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 3671-3676

**Pepe A., West Stephen C. (2014a)** MUS81-EME2 promotes replication fork restart. *Cell Reports* **7**, 1048-1055

**Pepe A., West S. C. (2014b)** Substrate specificity of the MUS81-EME2 structure selective endonuclease. *Nucleic Acids Research* **42**, 3833-3845

**Pepin J., Milord F. (1991)** African trypanosomiasis and drug-induced encephalopathy - risk factors and pathogenesis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **85**, 222-224

**Petermann E., Luis Orta M., Issaeva N., Schultz N., Helleday T. (2010)** Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. *Molecular Cell* **37**, 492-502

**Pinter S. F., Aubert S. D., Zakian V. A. (2008)** The *Schizosaccharomyces pombe* Pfh1p DNA helicase is essential for the maintenance of nuclear and mitochondrial DNA. *Molecular and Cellular Biology* **28**, 6594-6608

**Pluciennik A., Iyer R. R., Napierala M., Larson J. E., Filutowicz M., Wells R. D. (2002)** Long CTG-CAG repeats from myotonic dystrophy are preferred sites for intermolecular recombination. *Journal of Biological Chemistry* **277**, 34074-34086

**Polo S. E., Jackson S. P. (2011)** Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. *Genes & Development* **25**, 409-433

**Povelones M. L., Gluenz E., Dembek M., Gull K., Rudenko G. (2012)** Histone H1 plays a role in heterochromatin formation and VSG expression site silencing in *Trypanosoma brucei*. *PLoS Pathogens* **8**

**Prince P. R., Emond M. J., Monnat R. J. (2001)** Loss of Werner syndrome protein function promotes aberrant mitotic recombination. *Genes & Development* **15**, 933-938

**Proudfoot C., McCulloch R. (2005)** Distinct roles for two RAD51-related genes in *Trypanosoma brucei* antigenic variation. *Nucleic Acids Research* **33**, 6906-6919

**Proudfoot C., McCulloch R. (2006)** *Trypanosoma brucei* DMC1 does not act in DNA recombination, repair or antigenic variation in bloodstream stage cells. *Molecular and Biochemical Parasitology* **145**, 245-253

**Prucca C. G., Rivero F. D., Lujan H. D. (2011)** Regulation of antigenic variation in *Giardia lamblia*. *Annual Review of Microbiology* **65**, 611-630

**Rasband W. S.** <http://imagej.nih.gov/ij/> Accessed November 2014



- Raynard S., Bussen W., Sung P. (2006)** A double Holliday junction dissolvosome comprising BLM, topoisomerase III alpha, and BLAP75. *Journal of Biological Chemistry* **281**, 13861-13864
- Raynard S., Zhao W., Bussen W., Lu L., Ding Y.-Y., Busygina V., Meetei A. R., Sung P. (2008)** Functional role of BLAP75 in BLM-topoisomerase III alpha-dependent Holliday junction processing. *Journal of Biological Chemistry* **283**, 15701-15708
- Redmond S., Vadivelu J., Field M. C. (2003)** RNAit: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **128**, 115-118
- Redmond S V. J., Field MC (2003)** RNAit: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **128**, 115-118
- Redon C., Pilch D. R., Rogakou E. P., Orr A. H., Lowndes N. F., Bonner W. M. (2003)** Yeast histone 2A serine 129 is essential for the efficient repair of checkpoint-blind DNA damage. *EMBO Reports* **4**, 678-684
- Regairaz M., Zhang Y.-W., Fu H., Agama K. K., Tata N., Agrawal S., Aladjem M. I., Pommier Y. (2011a)** Mus81-mediated DNA cleavage resolves replication forks stalled by topoisomerase I-DNA complexes. *Journal of Cell Biology* **195**, 739-749
- Regairaz M., Zhang Y. W., Fu H., Agama K. K., Tata N., Agrawal S., Aladjem M. I., Pommier Y. (2011b)** Mus81-mediated DNA cleavage resolves replication forks stalled by topoisomerase I-DNA complexes. *Journal of Cell Biology* **195**, 739-749
- Reiter H., Kelley P., Milewski M (1972)** Mode of action of phleomycin on *Bacillus subtilis*. *Journal of Bacteriology* **111**, 586-6
- Reuner B., Vassella E., Yutzy B., Boshart M. (1997)** Cell density triggers slender to stumpy differentiation of *Trypanosoma brucei* bloodstream forms in culture. *Molecular and Biochemical Parasitology* **90**, 269-280
- Reynolds D., Cliffe L., Forstner K. U., Hon C. C., Siegel T. N., Sabatini R. (2014)** Regulation of transcription termination by glucosylated hydroxymethyluracil, base J, in *Leishmania major* and *Trypanosoma brucei*. *Nucleic Acids Research* **42**, 9717-9729
- Ribeyre C., Lopes J., Boule J.-B., Piazza A., Guedin A., Zakian V. A., Mergny J.L., Nicolas A. (2009)** The Yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences *in vivo*. *PLoS Genetics* **5**
- Ricchetti M., Dujon B., Fairhead C. (2003)** Distance from the chromosome end determines the efficiency of double strand break repair in subtelomeres of haploid yeast. *Journal of Molecular Biology* **328**, 847-862

- Rico E., Rojas F., Mony B. M., Szoor B., MacGregor P., Matthews K. R. (2013) Bloodstream form pre-adaptation to the tsetse fly in *Trypanosoma brucei*. *Frontiers in Cellular and Infection Microbiology* 3
- Rizzo A., Salvati E., Porru M., D'Angelo C., Stevens M. F., D'Incalci M., Leonetti C., Gilson E., Zupi G., Biroccio A. (2009) Stabilization of quadruplex DNA perturbs telomere replication leading to the activation of an ATR-dependent ATM signaling pathway. *Nucleic Acids Research* 37, 5353-5364
- Robertson A. B., Klungland A., Rognes T., Leiros I. (2009) DNA repair in mammalian cells: base excision repair: the long and short of it. *Cellular and Molecular Life Sciences* 66, 981-993
- Robinson N. P., Burman N., Melville S. E., Barry J. D. (1999) Predominance of duplicative VSG gene conversion in antigenic variation in African trypanosomes. *Molecular and Cellular Biology* 19, 5839-5846
- Rogakou E. P., Pilch D. R., Orr A. H., Ivanova V. S., Bonner W. M. (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *Journal of Biological Chemistry* 273, 5858-5868
- Rudd S. G., Glover L., Jozwiakowski S. K., Horn D., Doherty A. J. (2013) PPL2 translesion polymerase is essential for the completion of chromosomal DNA replication in the african trypanosome. *Molecular Cell* 52, 554-565
- Rudenko G., Chaves I., Dirks-Mulder A., Borst P. (1998) Selection for activation of a new variant surface glycoprotein gene expression site in *Trypanosoma brucei* can result in deletion of the old one. *Molecular and Biochemical Parasitology* 95, 97-109
- Rudenko G., McCulloch R., DirksMulder A., Borst P. (1996) Telomere exchange can be an important mechanism of variant surface glycoprotein gene switching in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 80, 65-75
- Sacconi S., Lemmers R., Balog J., van der Vliet P. J., Lahaut P., van Nieuwenhuizen M. P., Straasheijm K. R., Debipersad R. D., Vos-Versteeg M., Salviati L., Casarin A., Pegoraro E., Tawil R., Bakker E., Tapscott S. J., Desnuelle C., van der Maarel S. M. (2013) The FSHD2 Gene SMCHD1 Is a Modifier of Disease Severity in Families Affected by FSHD1. *American Journal of Human Genetics* 93, 744-751
- Saini N., Ramakrishnan S., Elango R., Ayyar S., Zhang Y., Deem A., Ira G., Haber J. E., Lobachev K. S., Malkova A. (2013) Migrating bubble during break-induced replication drives conservative DNA synthesis. *Nature* 502, 389-392
- San Filippo J., Sung P., Klein H. (2008) Mechanism of eukaryotic homologous recombination. *Annual Review of Biochemistry* 77, 229-257
- Schulz V. P., Zakian V. A. (1994) The *Saccharomyces* Pif1 DNA helicase inhibits telomere elongation and *de-novo* telomere formation. *Cell* 76, 145-155

- Shah J. S., Young J. R., Kimmel B. E., Iams K. P., Williams R. O. (1987) The 5' flanking sequence of a *Trypanosoma brucei* variable surface glycoprotein gene. *Molecular and Biochemical Parasitology* **24**, 163-174
- Shapiro S. Z., Naessens J., Liesegang B., Moloo S. K., Magondou J. (1984) Analysis by flow-cytometry of DNA synthesis during the life cycle of the African trypanosomes. *Acta Tropica* **41**, 313-323
- Sharma S., Brosh R. M., Jr. (2007) Human RECQ1 is a DNA damage responsive protein required for genotoxic stress resistance and suppression of sister chromatid exchanges. *PLoS ONE* **2**
- Shedden K., Vrughe D. T., Rudenko G. (2004) Bloodstream form-specific up-regulation of silent VSG expression sites and procyclin in *Trypanosoma brucei* after inhibition of DNA synthesis or DNA damage. *Journal of Biological Chemistry* **279**, 13363-13374
- Shi H., Tschudi C., Ullu E. (2006) An unusual Dicer-like1 protein fuels the RNA interference pathway in *Trypanosoma brucei*. *RNA* **12**, 2063-2072
- Shi H. F., Ullu E., Tschudi C. (2004) Function of the trypanosome argonaute 1 protein in RNA interference requires the N-terminal RGG domain and arginine 735 in the Piwi domain. *Journal of Biological Chemistry* **279**, 49889-49893
- Shiflett A. M., Bishop J. R., Pahwa A., Hajduk S. L. (2005) Human high density lipoproteins are platforms for the assembly of multi-component innate immune complexes. *Journal of Biological Chemistry* **280**, 32578-32585
- Sidorova J. M., Kehrli K., Mao F., Monnat R., Jr. (2013) Distinct functions of human RECQ helicases WRN and BLM in replication fork recovery and progression after hydroxyurea-induced stalling. *DNA Repair* **12**, 128-139
- Siegel T. N., Hekstra D. R., Kemp L. E., Figueiredo L. M., Lowell J. E., Fenyo D., Wang X., Dewell S., Cross G. A. M. (2009) Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes & Development* **23**, 1063-1076
- Simpson A. G. B., Stevens J. R., Lukeš J. (2006) The evolution and diversity of kinetoplastid flagellates. *Trends in Parasitology* **22**, 168-174
- Singh D. K., Karmakar P., Aamann M., Schurman S. H., May A., Croteau D. L., Burks L., Plon S. E., Bohr V. A. (2010) The involvement of human RECQL4 in DNA double-strand break repair. *Aging Cell* **9**, 358-371
- Singh T. R., Ali A. M., Busygina V., Raynard S., Fan Q., Du C.-h., Andreassen P. R., Sung P., Meetei A. R. (2008) BLAP18/RMI2, a novel OB-fold-containing protein, is an essential component of the Bloom helicase-double Holliday junction dissolvosome. *Genes & Development* **22**, 2856-2868
- Smith G. C. M., Jackson S. P. (1999) The DNA-dependent protein kinase. *Genes & Development* **13**, 916-934

- Songa E. B., Hamers R., Rickman R., Nantulya V. M., Mulla A. F., Magnus E. (1991) Evidence for widespread asymptomatic *Trypanosoma rhodesiense* human infection in the Luangwa Valley (Zambia). *Tropical Medicine and Parasitology* **42**, 389-393
- Štafa A., Donnianni R. A., Timashev L. A., Lam A. F., Symington L. S. (2014) Template switching during break-induced replication is promoted by the Mph1 helicase in *Saccharomyces cerevisiae*. *Genetics* **196**, 1017-1028
- Stanne T. M., Rudenko G. (2010) Active VSG Expression Sites in *Trypanosoma brucei* Are Depleted of Nucleosomes. *Eukaryotic Cell* **9**, 136-147
- Stavnezer J., Guikema J. E. J., Schrader C. E. (2008) Mechanism and regulation of class switch recombination. In *Annual Review of Immunology* **26**, 261-292.
- Stewart E., Chapman C. R., Al-Khodairy F., Carr A. M., Enoch T. (1997) rqh1+, a fission yeast gene related to the Bloom's and Werner's syndrome genes, is required for reversible S phase arrest. *EMBO Journal* **16**, 2682-2692
- Stewart J. A., Chaiken M. F., Wang F., Price C. M. (2012) Maintaining the end: Roles of telomere proteins in end-protection, telomere replication and length regulation. *Mutation Research* **730**, 12-19
- Stucki M., Clapperton J. A., Mohammad D., Yaffe M. B., Smerdon S. J., Jackson S. P. (2005) MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. *Cell* **123**, 1213-1226
- Stucki M., Jackson S. P. (2006) gamma H2AX and MDC1: Anchoring the DNA-damage-response machinery to broken chromosomes. *DNA Repair* **5**, 534-543
- Sugiyama T., Kowalczykowski S. C. (2002) Rad52 protein associates with replication protein A (RPA)-single-stranded DNA to accelerate Rad51-mediated displacement of RPA and presynaptic complex formation. *Journal of Biological Chemistry* **277**, 31663-31672
- Sugiyama T., Zaitseva E. M., Kowalczykowski S. C. (1997) A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the *Saccharomyces cerevisiae* Rad51 protein. *Journal of Biological Chemistry* **272**, 7940-7945
- Sung P. (1997) Communication - Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase. *Journal of Biological Chemistry* **272**, 28194-28197
- Sung P., Klein H. (2006) Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nature Reviews Molecular Cell Biology* **7**, 739-750
- Symington L. S., Gautier J. (2011) Double-strand break end resection and repair pathway choice. *Annu Rev Genet* **45**, 247-271

- Tadokoro T., Ramamoorthy M., Popuri V., May A., Tian J., Sykora P., Rybanska I., Wilson D. M., III, Croteau D. L., Bohr V. A. (2012) Human RECQL5 participates in the removal of endogenous DNA damage. *Molecular Biology of the Cell* **23**, 4273-4285
- Takata M., Sasaki M. S., Sonoda E., Morrison C., Hashimoto M., Utsumi H., Yamaguchi-Iwai Y., Shinohara A., Takeda S. (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO Journal* **17**, 5497-5508
- Takeuchi Y., Horiuchi T., Kobayashi T. (2003) Transcription-dependent recombination and the role of fork collision in yeast rDNA. *Genes & Development* **17**, 1497-1506
- Tanaka H., Ryu G. H., Seo Y. S., Tanaka K., Okayama H., MacNeill S. A., Yuasa Y. (2002) The fission yeast pfh1(+) gene encodes an essential 5' to 3' DNA helicase required for the completion of S-phase. *Nucleic Acids Research* **30**, 4728-4739
- Tarsounas M., Davies A. A., West S. C. (2004) RAD51 localization and activation following DNA damage. *Philosophical Transactions of the Royal Society B-Biological Sciences* **359**, 87-93
- Tercero J. A., Diffley J. F. X. (2001) Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint. *Nature* **412**, 553-557
- Tetley L., Turner C. M. R., Barry J. D., Crowe J. S., Vickerman K. (1987) Onset of expression of the variant surface glycoproteins of Trypanosoma brucei in the tsetse-fly studied using immunoelectron microscopy. *Journal of Cell Science* **87**, 363-372
- Thangavel S., Berti M., Levikova M., Pinto C., Gomathinayagam S., Vujanovic M., Zellweger R., Moore H., Lee E. H., Henclickson E. A., Cejka P., Stewart S., Lopes M., Vindigni A. (2015) DNA2 drives processing and restart of reversed replication forks in human cells. *Journal of Cell Biology* **208**, 545-562
- Thon G., Baltz T., Eisen H. (1989) Antigenic diversity by the recombination of pseudogenes. *Genes & Development* **3**, 1247-1254
- Thon G., Baltz T., Giroud C., Eisen H. (1990) Trypanosome variable surface glycoproteins - composite genes and order of expression. *Genes & Development* **4**, 1374-1383
- Thorslund T., McIlwraith M. J., Compton S. A., Lekomtsev S., Petronczki M., Griffith J. D., West S. C. (2010) The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. *Nature Structural & Molecular Biology* **17**, 1263-1265

- Timmers H. T. M., Delange T., Kooter J. M., Borst P. (1987) Coincident multiple activations of the same surface-antigen gene in *Trypanosoma brucei*. *Journal of Molecular Biology* **194**, 81-90
- Timmins J., McSweeney S. (2006) XPB: an essential helicase involved in both transcription and repair of DNA. *Molecular Cell* **22**, 149-150
- Trask B. J., Friedman C., Martin-Gallardo A., Rowen L., Akinbami C., Blankenship J., Collins C., Giorgi D., Iadonato S., Johnson F., Kuo W. L., Massa H., Morrish T., Naylor S., Nguyen O. T. H., Rouquier S., Smith T., Wong D. J., Youngblom J., van den Engh G. (1998) Members of the olfactory receptor gene family are contained in large blocks of DNA duplicated polymorphically near the ends of human chromosomes. *Human Molecular Genetics* **7**, 13-26
- Traverso G., Bettgowda C., Kraus J. r., Speicher M. R., Kinzler K. W., Vogelstein B., Lengauer C. (2003) Hyper-recombination and genetic instability in BLM-deficient epithelial cells. *Cancer Research* **63**, 8578-8581
- Trenaman A., Hartley C., Prorocic M., Passos-Silva D. G., van den Hoek M., Nechyporuk-Zloy V., Machado C. R., McCulloch R. (2013) *Trypanosoma brucei* BRCA2 acts in a life cycle-specific genome stability process and dictates BRC repeat number-dependent RAD51 subnuclear dynamics. *Nucleic Acids Research* **41**, 943-960
- Tuduri S., Crabbe L., Conti C., Tourriere H., Holtgreve-Grez H., Jauch A., Pantesco V., De Vos J., Thomas A., Theillet C., Pommier Y., Tazi J., Coquelle A., Pasero P. (2009) Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nature Cell Biology* **11**, 1315-1324
- Turner C. M. R. (1997) The rate of antigenic variation in fly-transmitted and syringe-passaged infections of *Trypanosoma brucei*. *FEMS Microbiology Letters* **153**, 227-231
- Turner C. M. R., Aslam N., Dye C. (1995) Replication, differentiation, growth and the virulence of *Trypanosoma brucei* infections. *Parasitology* **111**, 289-300
- Turner C. M. R., Barry J. D., Maudlin I., Vickerman K. (1988) An estimate of the size of the metacyclic variable antigen repertoire of *Trypanosoma brucei rhodesiense*. *Parasitology* **97**, 269-276
- Ui A., Satoh Y., Onoda F., Miyajima A., Seki M., Enomoto T. (2001) The N-terminal region of Sgs1, which interacts with Top3, is required for complementation of MMS sensitivity and suppression of hyper-recombination in sgs1 disruptants. *Molecular Genetics and Genomics* **265**, 837-850
- Umez K., Nakayama K., Nakayama H. (1990) *Escherichia coli* RecQ protein is a DNA helicase. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 5363-5367

- Unno J., Takagi M., Piao J., Sugimoto M., Honda F., Maeda D., Masutani M., Kiyono T., Watanabe F., Morio T., Teraoka H., Mizutani S. (2013) Artemis-dependent DNA double-strand break formation at stalled replication forks. *Cancer Science* **104**, 703-710
- Valdes J., Taylor M. C., Cross M. A., Ligtenberg M. J., Rudenko G., Borst P. (1996) The viral thymidine kinase gene as a tool for the study of mutagenesis in *Trypanosoma brucei*. *Nucleic Acids Research* **24**, 1809-1815
- Vandyck E., Foury F., Stillman B., Brill S. J. (1992) A single-stranded DNA binding protein required for mitochondrial DNA replication in *Saccharomyces cerevisiae* is homologous to *Escherichia coli* SSB. *EMBO Journal* **11**, 3421-3430
- Vanhamme L., Poelvoorde P., Pays A., Tebabi P., Xong H. V., Pays E. (2000) Differential RNA elongation controls the variant surface glycoprotein gene expression sites of *Trypanosoma brucei*. *Molecular Microbiology* **36**, 328-340
- Vassella E., Reuner B., Yutzy B., Boshart M. (1997) Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. *Journal of Cell Science* **110**, 2661-2671
- Veaute X., Jeusset J., Soustelle C., Kowalczykowski S. C., Le Cam E., Fabre F. (2003) The Srs2 helicase prevents recombination by disrupting Rad51 nucleoprotein filaments. *Nature* **423**, 309-312
- Viguera E., Canceill D., Ehrlich S. (2001) Replication slippage involves DNA polymerase pausing and dissociation. *EMBO Journal* **20**, 2587-2595
- Wagner M., Price G., Rothstein R. (2006) The absence of Top3 reveals an interaction between the Sgs1 and Pif1 DNA Helicases in *Saccharomyces cerevisiae*. *Genetics* **174**, 555-573
- Wallis J. W., Chrebet G., Brodsky G., Rolfe M., Rothstein R. (1989) A hyper-recombination mutation in *S. cerevisiae* identifies a novel eukaryotic topoisomerase. *Cell* **58**, 409-419
- Wang J., Englund P. T., Jensen R. E. (2012) TbPIF8, a *Trypanosoma brucei* protein related to the yeast Pif1 helicase, is essential for cell viability and mitochondrial genome maintenance. *Molecular Microbiology* **83**, 471-485
- Wang T. F., Kung W. M. (2002) Supercomplex formation between Mlh1-Mlh3 and Sgs1-Top3 heterocomplexes in meiotic yeast cells. *Biochemical and Biophysical Research Communications* **296**, 949-953
- Wang W., Seki M., Narita Y., Nakagawa T., Yoshimura A., Otsuki M., Kawabe Y.-i., Tada S., Yagi H., Ishii Y., Enomoto T. (2003) Functional relation among RecQ family helicases RecQL1, RecQL5, and BLM in Cell Growth and sister chromatid exchange formation. *Molecular and Cellular Biology* **23**, 3527-3535

- Wang W. S., Seki M., Narita Y., Sonoda E., Takeda S., Yamada K., Masuko T., Katada T., Enomoto T. (2000) Possible association of BLM in decreasing DNA double strand breaks during DNA replication. *EMBO Journal* **19**, 3428-3435
- Weitzmann M. N., Woodford K. J., Usdin K. (1996) The development and use of a DNA polymerase arrest assay for the evaluation of parameters affecting intrastrand tetraplex formation. *Journal of Biological Chemistry* **271**, 20958-20964
- Weller G. R., Kysela B., Roy R., Tonkin L. M., Scanlan E., Della M., Devine S. K., Day J. P., Wilkinson A., di Fagagna F. d. A., Devine K. M., Bowater R. P., Jeggo P. A., Jackson S. P., Doherty A. J. (2002) Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* **297**, 1686-1689
- West S. C., Blanco M. G., Chan Y. W., Matos J., Sarbajna S., Wyatt H. D. (2015) Resolution of Recombination Intermediates: Mechanisms and Regulation. *Cold Spring Harbor symposia on quantitative biology*
- Wickstead B., Ersfeld K., Gull K. (2003) The frequency of gene targeting in *Trypanosoma brucei* is independent of target site copy number. *Nucleic Acids Research* **31**, 3993-4000
- Wickstead B., Ersfeld K., Gull K. (2004) The small chromosomes of *Trypanosoma brucei* involved in antigenic variation are constructed around repetitive palindromes. *Genome Research* **14**, 1014-1024
- Wilkinson S. R., Kelly J. M. (2009) Trypanocidal drugs: mechanisms, resistance and new targets. *Expert Reviews in Molecular Medicine* **11**
- Wilson M. A., Kwon Y., Xu Y., Chung W.-H., Chi P., Niu H., Mayle R., Chen X., Malkova A., Sung P., Ira G. (2013) Pif1 helicase and Pol[delta] promote recombination-coupled DNA synthesis via bubble migration. *Nature* **502**, 393-396
- Woodward R., Gull K. (1990) Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*. *Journal of Cell Science* **95**, 49-57
- World Health Organisation  
<http://www.who.int/mediacentre/factsheets/fs259/en/> Accessed 15/04/2015
- Wright J. R., Siegel T. N., Cross G. A. M. (2010) Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **172**, 141-144
- Wu L., Davies S. L., Levitt N. C., Hickson I. D. (2001) Potential role for the BLM helicase in recombinational repair via a conserved interaction with RAD51. *Journal of Biological Chemistry* **276**, 19375-19381



- Wu L., Davies S. L., North P. S., Goulaouic H., Riou J. F., Turley H., Gatter K. C., Hickson I. D. (2000) The Bloom's syndrome gene product interacts with topoisomerase III. *Journal of Biological Chemistry* **275**, 9636-9644
- Wu L., Hickson I. D. (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* **426**, 870-874
- Wyatt M. D., Pittman D. L. (2006) Methylating agents and DNA repair responses: Methylated bases and sources of strand breaks. *Chemical Research in Toxicology* **19**, 1580-1594
- Xu X., Rochette P. J., Feyissa E. A., Su T. V., Liu Y. (2009a) MCM10 mediates RECQ4 association with MCM2-7 helicase complex during DNA replication. *EMBO Journal* **28**, 3005-3014
- Xu X. H., Rochette P. J., Feyissa E. A., Su T. V., Liu Y. L. (2009b) MCM10 mediates RECQ4 association with MCM2-7 helicase complex during DNA replication. *EMBO Journal* **28**, 3005-3014
- Yamagata K., Kato J., Shimamoto A., Goto M., Furuichi Y., Ikeda H. (1998) Bloom's and Werner's syndrome genes suppress hyperrecombination in yeast sgs1 mutant: implication for genomic instability in human diseases. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 8733-8738
- Yang J., Bachrati C. Z., Ou J., Hickson I. D., Brown G. W. (2010) Human topoisomerase IIIalpha is a single-stranded DNA decatenase that is stimulated by BLM and RMI1. *Journal of Biological Chemistry* **285**, 21426-21436
- Yin J. H., Sobeck A., Xu C., Meetei A. R., Hoatlin M., Li L., Wang W. D. (2005) BLAP75, an essential component of Bloom's syndrome protein complexes that maintain genome integrity. *EMBO Journal* **24**, 1465-1476
- Young R., Taylor J. E., Kurioka A., Becker M., Louis E. J., Rudenko G. (2008) Isolation and analysis of the genetic diversity of repertoires of VSG expression site containing telomeres from *Trypanosoma brucei gambiense*, *T. b. brucei* and *T. equiperdum*. *BMC Genomics* **9**
- Zeman M. K., Cimprich K. A. (2014) Causes and consequences of replication stress. *Nature Cell Biology* **16**, 2-9
- Zhang J. R., Hardham J. M., Barbour A. G., Norris S. J. (1997) Antigenic variation in Lyme disease borreliae by promiscuous recombination of VMP-like sequence cassettes. *Cell* **89**, 275-285
- Zhang Y., Rohde L. H., Wu H. (2009) Involvement of nucleotide excision and mismatch repair mechanisms in double strand break repair. *Current Genomics* **10**, 250-258
- Zhou J. Q., Monson E. K., Teng S. C., Schulz V. P., Zakian V. A. (2000) Pif1p helicase, a catalytic inhibitor of telomerase in yeast. *Science* **289**, 771-774

- Zhou J. Q., Qi H., Schulz V. P., Mateyak M. K., Monson E. K., Zakian V. A. (2002) *Schizosaccharomyces pombe pfh1+* encodes an essential 5' to 3' DNA helicase that is a member of the PIF1 subfamily of DNA helicases. *Molecular Biology of The Cell* **13**, 2180-2191
- Zhu Z., Chung W.-H., Shim E. Y., Lee S. E., Ira G. (2008) Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. *Cell* **134**, 981-994